International Energy Agency
Energy Technology Initiative on
Demand Side Management Technologies and Programmes

Final Status Report: U.S.A.

Task 24 – Phase II

Behaviour Change in DSM: Helping the Behaviour Changers

December 2018 *Dr. Sea Rotmann and Kira Ashby*

With many thanks to Dr. Beth Karlin (SCI), CEE staff Veronica Alix and Jennifer Smith, and our 11 CEE member sponsors who made US participation in Task 24 possible

Contents

Acronyms and Abbreviations

BCF	Behaviour Changer Framework
BRO	Behavioural Retro-Commissioning
CEE	Consortium for Energy Efficiency
CIA	Collective Impact Approach
CPUC	California Public Utility Commission
DOE	US Department of Energy
DSM	Demand-Side Management
EA	Electricity Authority
EECA	Energy Efficiency and Conservation Authority
EED	Energy Efficiency Directive
EEffG	Energy Efficiency Act
EEOSs	Energy Efficiency Obligation Schemes
EM&V	Evaluation, Measurement & Verification
ESCo	Energy Service Company
EVT	Efficiency Vermont
ExCo	Executive Committee (governing body of IEA DSM Technology Collaboration Programme)
HERs	Home Energy Reports
ICT	Information and Communication Technology
IOUs	Investor Owned Utilities
M&V	Monitoring & Verification
MBIE	Ministry of Business, Innovation, and Employment
MSs	Member States
NEBs	Non-Energy Benefits
NEIs	Non-Energy Impacts
NMEC	Normalized Metering Energy Consumption
NRDC	Natural Resources Defense Council
PAR	Participatory Action Research
PG&E	Pacific Gas & Electric Company
RCTs	Randomised Control Trials
RPP	Retail Products Platform
RRIM	Risk Reward Incentive Mechanism
SCI	See Change Institute
SEA	Sustainable Energy Advice Ltd
SEM	Strategic Energy Management
SMBs	Small and Medium Businesses
SMEs	Small and Medium Enterprises
ST	Subtask
TRC	Total Resource Cost Test

Executive Summary

The IEA Demand-Side Management (DSM) Task 24 aims at sharing knowledge between multiple stakeholder sectors and countries, and developing recommendations about the influence of behaviour change on effective implementation of energy-efficiency policies and programmes. The United States, in a joint collaboration between the Consortium for Energy Efficiency (CEE) and the US Department of Energy (DOE), joined Task 24 in Year 3 of Phase II. The 11 CEE member organisation sponsors selected the following overarching issue to focus on for US participation: *Behavioural programme evaluation methodologies, credibility, and persistence.*

The objective of this report is to examine how behaviour-based energy programmes are defined, implemented, and measured in North America (the US and Canada) and how this compares and contrasts with international approaches. This work is based on:

- An analysis of recent evaluation and behaviour change literature;
- An analysis of the public version of the CEE *Behavior Program Summary*;
- An analysis of Task 24 case studies collected from 2012-18;
- Two Task 24 workshops in the US in 2018;
- Interviews with CEE member sponsors, international experts and a former US Regulator from California; and
- Input from international experts during Task 24 sessions at two major international behaviour change conferences.

Behavioural Savings Potential and Related Challenges

Research suggests that energy-saving behaviours could deliver up to 20 percent reduction in greenhouse gas emissions from household energy usage (Dietz et al., 2009). In practice, however, human beings are much less predictable than energy-efficient equipment. The result is that a higher standard for evaluation is necessary to confidently determine savings for an energy efficiency effort that depends primarily on behaviour change as compared to a programme that relies on new equipment to achieve energy savings. As a result, the randomised control trial (RCT) became a gold standard for behavioural programmes early in the process of these approaches gaining acceptance in the energy efficiency industry. Yet not all programme approaches are evaluable via RCTs; thus, additional rigorous and credible evaluation approaches are needed to facilitate the implementation of additional programmes. Given the lack of accepted rigorous and credible evaluation approaches for behavioural approaches in the US, many energy efficiency efforts that would be considered behavioural in other countries are not labelled as such in the US, and it can therefore be difficult to demonstrate the value of, and gain approval for programmes that contain behavioural elements.

Regulatory Heterogeneity

Even within the US, there is a vast difference across states and jurisdictions in terms of the types of behavioural efforts that may be implemented. For example, in California, the initial definition allowed claimable savings for a subset of behavioural interventions that employed: 1) comparative energy usage, 2) randomised control trials (RCTs), and 3) ex-post measurement. In practice, the result has been that most behavioural programmes in California are Home Energy Reports (HERs), which are measured using RCTs, and typically generate around 1-2 percent energy savings. Elsewhere, somewhat broader regulatory definitions have allowed more flexible social science research and piloting. Most Task 24 countries' programmes also explored myriad behavioural strategies that have been used to successfully change behaviour across sectors, unlocking additional energy savings.

Barriers and Opportunities

Results from this US-based research identified several key barriers and programme opportunities. Barriers mentioned by CEE utility members often, but not always, converged around regulatory frameworks and included the difficulties of maintaining control groups, avoiding double counting (e.g. with widget-based programmes), and accounting for the wider error margins inherent in human behaviour. Opportunities included:

- Moving from an attribution to a contribution model (i.e. instead of attributing behaviour change to single actions, contribution models hope to uncover how each action contributes to change);
- Re-defining behavioural savings programmes;

- Increasing the number of evaluation strategies and methodologies that are considered rigorous and acceptable;
- Striving towards achieving sufficient confidence around behaviour for certain programmes to be able to "deem" the savings (i.e. assume a pre-approved quantity of energy savings, as is common in technology-based programmes); and
- Integrating international best practices.

Looking to Peers Abroad

An examination of both North American (US and Canadian) and international approaches to defining, implementing, and evaluating behavioural programmes has made it quite evident that there are no silver bullet solutions. We also recognise that there are cultural and societal differences between the US and Canada and the European countries from which many of the Task 24 case studies were drawn. That said, the absence of such a rigorous regulatory structure abroad that determines which types of behavioural approaches may and may not be explored has resulted in a variety of rich examples of how behavioural approaches can be used to shift energy usage behaviours. Although these approaches would not necessarily be transferable to the US without modification, they provide a glimpse into what behavioural techniques may be appropriate for adaptation or testing in the US.

"Behavioural Tools and Processes" vs. "Behavioural Programmes"

In conclusion, we propose moving away from a specific definition of "behaviour" or "behaviour programmes," and moving towards the concept of behaviour as a process or a set of tools to be utilised across energy efficiency programme types (Task 24 Project Partner, See Change Institute (SCI), outlined such a process in Workshop 2 in October 2018). Programmes following this framework will be better positioned to measure results, capture resulting energy savings, and demonstrate credibility. This framework helps ensure that both process and outcomes are measured, to allow programmes to be improved over time, thereby building the overall case for claiming subsequent savings. This framework is grounded in robust social science theory to ensure that resulting trials, tests, pilots or programmes are better able to be evaluated, therefore providing greater opportunity to replicate results through similar programmes. We will test such a process in the follow-up research collaboration from Task 24 – a new IEA DSM Task on hard-to-reach (HTR) energy users, which was inspired by this US collaboration.

Introduction

The IEA Demand-Side Management Task 24 aims at sharing knowledge between multiple stakeholder sectors and developing recommendations about the influence of behaviour change in the effective implementation of energy efficiency policies[1] and programmes. After a period of developing a scientific framework and collecting practical cases (Phase I), Task 24 has now completed Phase II, engaging actual "*Behaviour Changers*" in real life interventions, supporting them with evidence-based scientific approaches and practical case study comparisons from various countries along the way.

Task 24 and North America (US and Canada)

The US, via national expert CEE, has formally participated in Task 24 since the beginning of 2018, joining Year 3 of Phase II. Since 2014, Project Partner See Change Institute (SCI) has provided in-kind support on the topic of evaluating behavioural interventions to the Task. Other participants in Phase II of Task 24 were New Zealand, the Netherlands, Austria, Ireland and Sweden. The North American (US and Canadian) contribution was supported by CEE member sponsors: eight US and two Canadian utilities, along with two government agencies (the US Department of Energy and Natural Resources Canada). The Task's National Expert for the US, Kira Ashby, heads the behavioural programme at CEE. This report will concentrate on the North American-specific (US and Canada) theme of "**Evaluating behavioural programmes: methodology, credibility, and persistence**." For an in-depth description of the second phase of Task 24, the overall approach, and a detailed overview of the Subtasks, please refer to Appendix 1.

Background on the Consortium for Energy Efficiency (CEE)

The Consortium for Energy Efficiency (CEE) acted as the participating agent for US participation in Phase II of Task 24 in 2018. CEE is an award-winning consortium of efficiency programme administrators from the United States and Canada that unifies programme approaches across jurisdictions to increase impact in fragmented markets. By joining forces at CEE, individual electric and gas efficiency programmes are able to partner not only with each other, but also with other industries, trade associations, and government agencies. Working together, administrators leverage the effect of their ratepayer funding, exchange information on successful practices and, by doing so, achieve greater energy efficiency for the public good. CEE is a non-profit, non-governmental organisation, and its members include 83 organisations across the US and Canada.

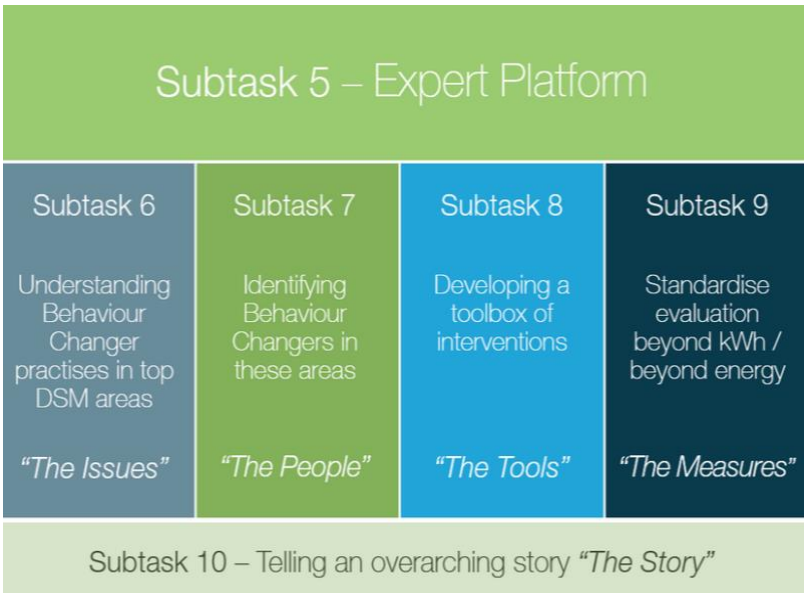

Figure 1. Task 24 Phase II Subtask overview

[1] See Task 24 Policy Brief: http://www.ieadsm.org/wp/files/task24policybrief.pdf

Benefits of an IEA research collaboration on behaviour

In a review of 15 years of energy social science research, Sovacool (2014) noted the value of research collaborations that span both countries and disciplines: "*Comparative studies, by producing more data than a single case, can more rigorously generate and test hypotheses across multiple areas, resulting in stronger evidence through a convergence of findings, and a wider applicability of results. Moreover, when researchers from different backgrounds are incentivised to conduct collaborative and cross-national projects, they can capitalise on their strengths and offset potential weaknesses*."

Hantrais (1995) observed that cross-country comparisons can add depth to our collective knowledge base: "*Comparisons can lead to fresh, exciting insights and a deeper understanding of issues that are of central concern in different countries. They can lead to the identification of gaps in knowledge and may point to possible directions that could be followed and about which the researcher may not previously have been aware. They may also help to sharpen the focus of analysis of the subject under study by suggesting new perspectives*."

Most analyses of behavioural interventions do not explicitly focus on country comparisons. This gap heavily influenced the establishment of different IEA research contracts between different countries. In Subtask 2 (Phase I), we focused explicitly on such cultural idiosyncrasies. Cultural differences and their origins (cultural traits or a particular cultural characteristic) do impact the meaningfulness and efficacy of generic policy recommendations for *Behaviour Changers*. Identifying various cultural contexts and designing and testing a toolbox for behavioural interventions that works in many different countries, sectors, and DSM issues, were major objectives of this Task. Policy briefings specific to the participating countries' policy makers were developed, including for the US. On the top US issue of *evaluating behavioural programmes* we received input from Task 24 experts and undertook a cross-country comparison between the US and Canada, Austria, Sweden and New Zealand.

The added value to having an IEA Expert Platform (Subtask 5) is a highly experienced global network of *Behaviour Changers* in many different countries, sectors, disciplines and industries. These individuals can all offer different insights, learnings and perspectives, and many of them do so in-kind. We facilitated their collaboration with national *Behaviour Changers* by using and testing the *Collective Impact Approach* (CIA) for the first time in the energy system. The *Behaviour Changers* participating in this Task have assessed the effectiveness of this approach and the Task 24 Subtask 8 Toolbox for *Behaviour Changers* (Rotmann 2018a). The US national expert had a chance to collaborate in person with many of these Behaviour Changers during the BEHAVE conference in Switzerland in September 2018. Several of the experts who participated in Task 24 are already involved in a new IEA DSM Task on "Hard-to-Reach Energy Users", which was inspired by CEE sponsors' input on the challenges of successfully implementing and evaluating behavioural programmes targeting this important energy user group.

Behaviour Definitions

Task 24 Definitions

During the first international Task 24 workshop at Oxford University in October 2012 (Churchhouse, Mahoney & Rotmann 2012), it became apparent that we had to be very careful with the language and terminology used in this Task. Given that the language used around behaviour change across different research disciplines and sectors can vary widely, it is quite easy to confuse meanings and terminology. In order to clarify what language the Task was using, we had to create our own definitions for the main terms: *energy behaviour, behaviour change, Behaviour Changer, behavioural models, DSM, evaluation, monitoring, effectiveness, efficiency, investment versus habitual behaviours, outputs versus outcomes, single- and double-loop learning, and DSM tools and benchmarks* (found in Rotmann and Mourik, 2013). The most important definitions used are replicated below:

Energy behaviour refers to all human actions that affect the way that fuels - including electricity, natural gas, petroleum (gasoline), coal, etc. - are used to achieve desired services, including the acquisition or disposal of energy-related technologies and materials, the ways in which these are used, and the mental processes that govern these actions.

Behaviour Change in the context of this Task thus refers to any changes in said human actions which were directly or indirectly influenced by a variety of interventions (e.g. legislation, regulation, energy

efficiency programmes and technologies, incentives, subsidies, information campaigns, word-of-mouth etc.) aimed at fulfilling specific behaviour change outcomes. These outcomes can include any changes in energy efficiency, total energy consumption, energy technology uptake or demand-side management, but should be identified and specified by the *Behaviour Changer* designing the intervention for the purpose of outcome evaluation.

Behaviour Changer is a person or agency tasked with the goal of designing, implementing, evaluating, and/or disseminating interventions geared at changing energy *End User* behaviours. In this Task, we differentiate between five *Behaviour Changer* sectors: "the Decision-maker" (usually government on all levels), "the Provider" (usually energy- and energy technology-providing industry on all levels), "the Expert" (researchers and consultants from a multitude of disciplines, especially economics, psychology, sociology and engineering), "the Conscience" (the third sector including NGOs, community organisations, consumer groups, advocacy organisations, etc.), and "the Middle Actor" (usually goods and service providers and contractors in direct contact with the End Users).

Evaluation Definitions

Evaluation is a structured process of assessing the success of an intervention in meeting its goals and reflecting on the lessons learned during the intervention.

Monitoring refers to measuring progress and achievement, and whether the planned outputs and outcomes are produced.

Effectiveness refers to the extent to which an intervention reaches the intended goals.

Efficiency is usually measured in terms of cost-effectiveness, which compares the inputs in terms of staff time and financial resources with the outputs in terms of energy or cost savings achieved.

Outputs are direct and measurable products of an intervention.

Outcomes refer to the results and impact and/or improvements in the short, medium, and long-term.

Single-loop learning is about the effectiveness and/or efficiency of a technology, measure, instrument, arrangement, or intervention to achieve predefined goals.

Double-loop learning is process-oriented, focused on the how, when, where, how, how long, for whom and is about questioning goals and the prevailing norms and rules underlying these goals. In addition, double-loop learning is focused on interactions, the quality of participation, learning by doing and doing by learning, aligning expectations. In short, double-loop learning is about iterative governance of interventions.

A benchmark is a point of reference from which measurements may be made.

Methodology for developing this research topic

Prior to the first Task 24 Workshop in April 2018, CEE surveyed the sponsors and held in-depth discussions during the workshop. The sponsors proposed top issues to the following five topics:

1. **Evaluation** - "Beyond kWh" Survey Toolkit
2. **Evaluation** – Non-Energy Benefits
3. **Evaluation** - Methods, credibility, and persistence
4. **Small to Medium Businesses**
5. **Smart/Connected Home**

The following criteria were used to select the final topic:

1. **Whether the topic leverages IEA-DSM Task 24 assets**, particularly based on the Subtask 3 Evaluation, Measurement & Verification (EM&V) report (Mourik et al 2015).
2. **Whether the topic leverages CEE / sponsor assets**, such as the CEE *Behavior Program Summary*, which provides an overview of CEE members' behavioural programmes and has been published annually since 2010.

3. **Whether the topic leverages international knowledge/learnings**, from other countries - partly from Subtask 1, 2 and 6 case study analyses (with specific examples from Switzerland, Austria, New Zealand, the UK, and Ireland) and especially the Austrian Subtask 6 country case study (Kallsperger & Rotmann, 2018) which looked at regulatory barriers in EM&V.
4. **The topic's feasibility,** and whether it could realistically be addressed in the one year of US participation.

Given these criteria, the sponsors selected the following topic via online voting:

Evaluation - Methods, Credibility, and Persistence

In this research project, we surveyed evaluation techniques and approaches that have worked for different CEE sponsors, in other countries, and in different disciplines, including evidence for the credibility of behavioural programmes, and learnings from the persistence of behavioural savings after programmes end. This also included a preliminary assessment of methods for engaging hard-to-reach utility customers.

Goals of Data Collection

Goals of US Participation:

- Identify behavioural programme approaches from abroad that have (1) generated energy savings; (2) have been rigorously evaluated, and therefore may be more likely to be viewed as credible by US regulators; and (3) could be promising prospects for sponsors to pilot in the US and Canada.
- Identify evaluation approaches and techniques that have been effective at evaluating behavioural energy efficiency programmes abroad; consider which of these approaches may be the most compelling to regulators in the US and Canada.
- Learn whether other countries have studied the persistence of their behavioural programmes and if so, learn more about which behavioural programme type(s) have been studied and what the findings have been.
- Better understand which behavioural programmes abroad are designed for hard-to-reach energy users and in which countries these programmes are taking place, including details about the programme design and implementation.

Task 24 Goals:

- Add the US as another country participant and gain insights from the world's major economy, which also has a unique regulatory environment.
- Undertake an in-depth review on the chosen topic by US and Canadian sponsors.
- Add the US and Canadian insights into the Subtask 10 overarching story report.

In summary, the US contribution to Task 24 was shaped using the following steps:

Step	Deliverables	Method
1	Identification of the top DSM issues in US (ST 6 "The Issues")	Workshop 1, internal CEE member discussions, survey of CEE sponsors
2	Identification of the Behaviour Changers in the US; national & international expert network (ST 7 "The People")	Workshops 1 & 2, CEE members, SCI project partner, Task 24 expert network, BEHAVE and BECC Task 24 special sessions
3	Application and testing of Task 24 tools (ST 8 "The Tools")	Two Task 24 Workshops in US, BEHAVE conference 2018, BECC session 2018
4	Input for the final report (ST 10 "The US Report")	Stakeholder and expert surveys, CEE *Behavior Program Summary* analysis, Task 24 case study analysis, input and feedback from conference and workshop attendees, literature review, and interviews with CEE member organisations, IEA DSM Task 24 experts, and a former US Regulator.

Timeline

The CEE sponsors joined Year 3 of Phase II of Task 24, compressing the usual three-year timeline (see Figure 2). Project partner SCI has collaborated on evaluation methodology since 2014. Concurrently, a Subtask 11 case study was happening in the largest US health network in the Carolinas (see Cowan et al., 2017 and 2018).

Figure 2. Timeline of CEE / Task 24 Year 3 workflow.

Findings

The main findings from the CEE sponsors are structured into four main parts: 1) main DSM-issues in US and Canada; 2) top issue on evaluating behavioural programmes (methodologies, credibility, persistence, and hard-to-reach customers); 3) international insights and comparisons; and 4) remaining questions.

1) Overview of Main DSM Issues in the US and Canada: Workshop 1, April 10

National Potential for Energy Efficiency and Demand-Side Management

During the first Workshop, held in April 2018 (described in detail in the combined US workshop minutes), several barriers and market failures in the North American (US and Canadian) context were discussed. Before the Workshop, participating CEE member utilities completed a survey to help shed light on their main areas of interest, main challenges, and current behavioural approaches (see Figure 3).

The top behavioural issues of interest to US & Canadian Behaviour Changers

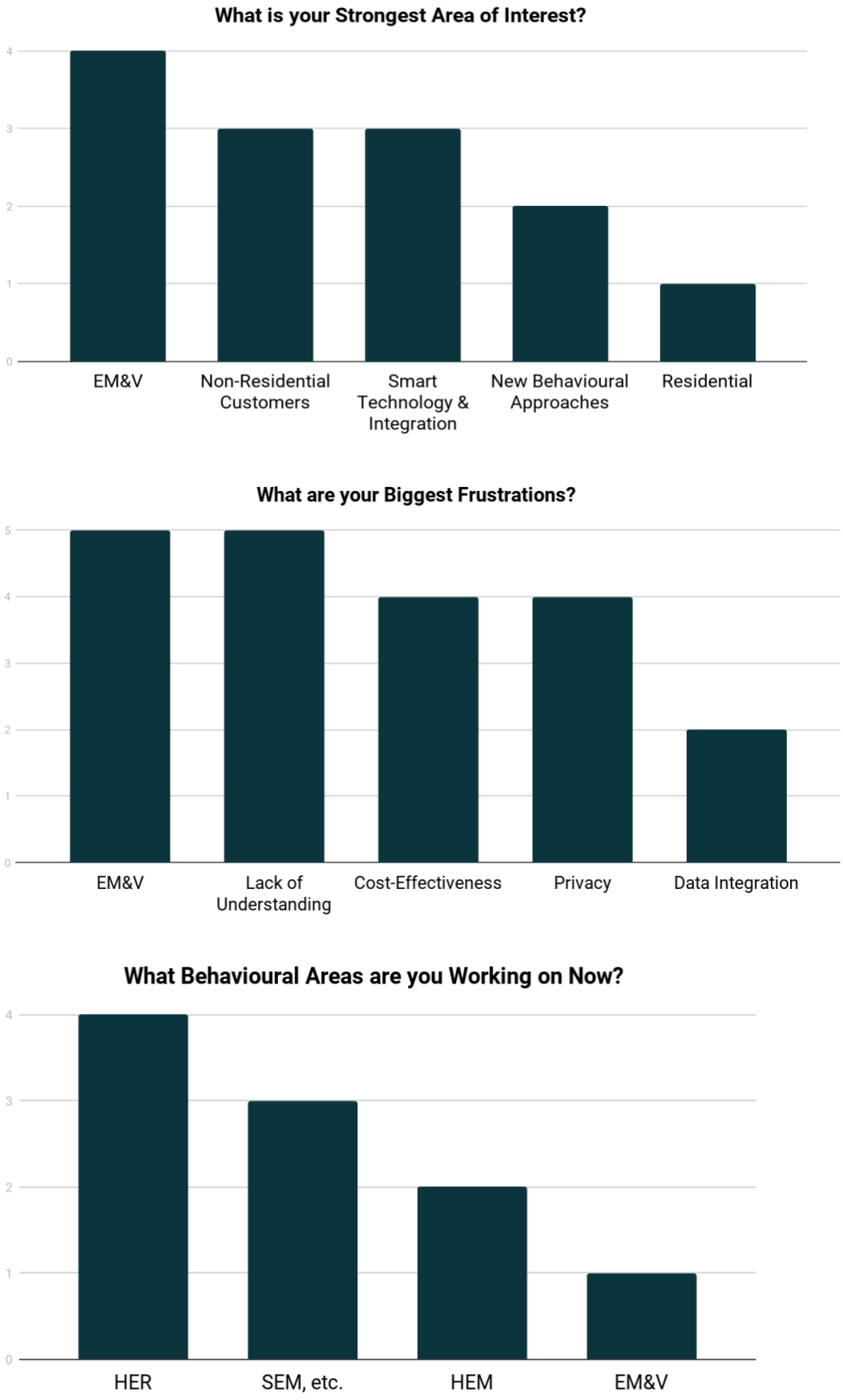

Figure 3. Summary of results from pre-workshop sponsor survey.

Using the Task 24 *Behaviour Changer Framework*

At the April 2018 Workshop, CEE sponsors and members engaged in a deep discussion around their most important issues related to behavioural interventions. These key topics were diverse, and spanned the residential and commercial sectors, touched on methodology and design (including evaluation), and included barriers such as the diverse regulatory environments that exist across the US and Canada. For the purposes of working through the Task 24 "Magic Carpet" exercise during Workshop 1, in which the different key stakeholders and their relationships are identified, CEE participants selected the topic of "Vampire stand-by power in the residential sector" (see Figure 4).

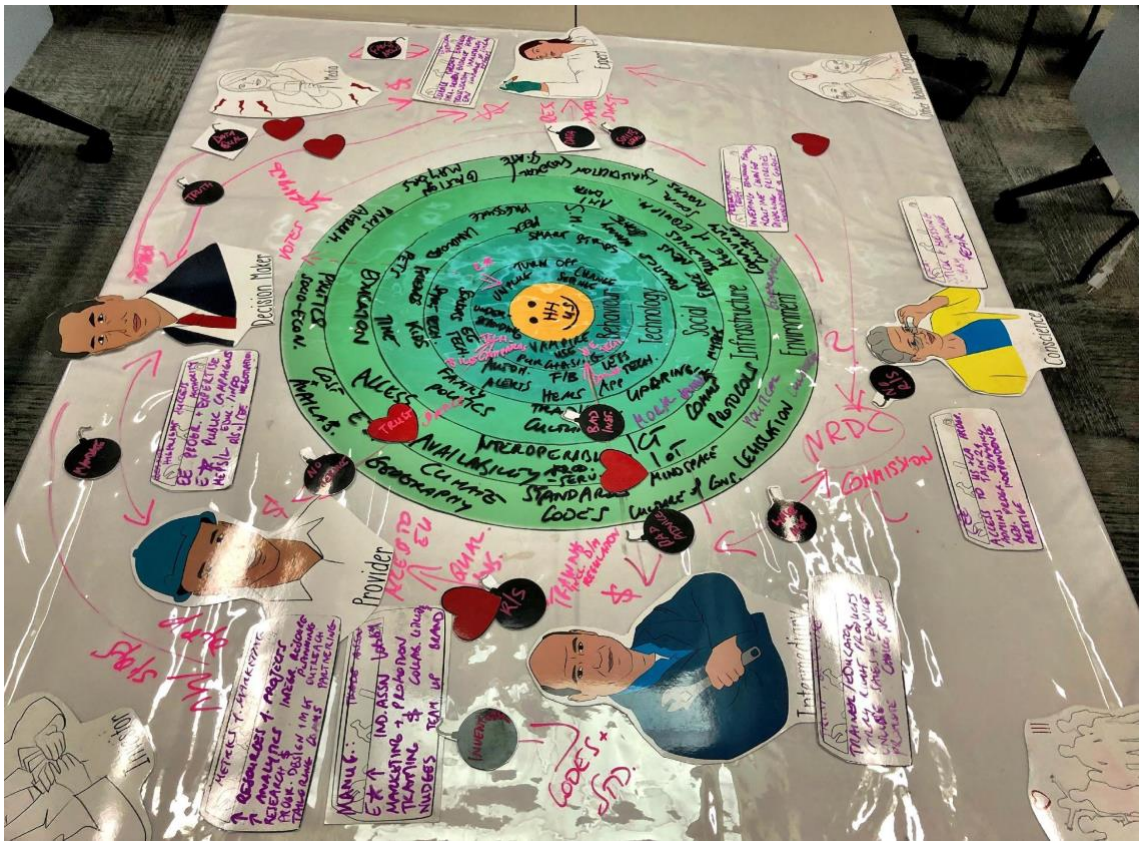

Figure 4. "Magic carpet" exercise, Workshop 1, April 10, 2018.

Behaviour Changers in the US

In Workshops 1 and 2, we brought together CEE sponsors and members, *Decision-makers* (US DOE), *Providers* (the 11 CEE sponsors), *Experts* (SCI, Task 24 and Jennifer Senick, Rutgers Center for Green Buildings) and the *Middle Actor* (CEE). During the "magic carpet" exercise in Workshop 1 (see below), it became clear that we were missing *the Conscience* (e.g. the Natural Resource Defence Council, also a CEE member). Other important *Decision-makers* that were identified as missing were the Regulators who determine which behavioural programmes can claim savings and how these savings should be evaluated. One former Regulator was later interviewed by the Task 24 Operating Agent to collect some of these insights for this report.

The 11 CEE sponsors represent all major US regions, in addition to several Canadian provinces, and include a Canadian Decision-maker (NRCan). Given the key role the sponsors played in the project, their representation of 24.6 million electric customers and 8.5 million natural gas customers throughout the US and Canada was essential to this report's findings.

The Behaviour Changers' main mandates, stakeholders, restrictions, and tools

The "magic carpet" exercise is a heuristic to help visualise the socio-ecosystem of any given behaviour we hope to change with a behavioural intervention. For this exercise during Workshop 1, we looked at the behaviour of **reducing vampire energy in residential households**. We collected the main mandates, stakeholders, restrictions and tools for each *Behaviour Changer* and End User (see Table 1 below). This exercise helps *Behaviour Changers* better understand the different contexts each actor operates within

and identify commonalities or additional tools that can help in designing behavioural interventions. Then, we looked at the End User context (for the chosen behaviour, see Table 2) and after that, we assessed the strength of relationships between each *Behaviour Changer* and their relationship with the End User (see Figure 4). Finally, we checked if there were special opportunities (designated by love hearts) or conflicts (bombs) to highlight in the system. This final step helps to identify potential End User pain points when designing an intervention, along with barriers and opportunities that can help or hinder progress. For a more detailed description of the process, see Rotmann (2016).

Table 1. Main mandates, stakeholders, restrictions, and tools of Behaviour Changers.

Behaviour Changer		**Mandate**	**Stakeholders**	**Restrictions**	**Tools**
End Users (Residential households)		• For low income end users: keep family warm/fed & bills paid • For middle income end users: keep bills paid and live life • For high income end users: be green	Family Pets Visitors	It is inconvenient to turn everything off/unplug, don't want to reset clocks, accidentally unplug the wrong thing, budget, family politics	Purchasing behaviour, habit change, educating and engaging family members, diverging priorities in the home
Decision-Makers	Board of Directors	Meet corporate goals, meet shareholder expectations, drive profit-related activities, company providing secure / reliable energy	Shareholders, fellow Board members, fellow employees, the organisation/ ratepayers/ customers	Regulation mandates, funding, risk tolerance/ aversion, no information / knowledge of behaviour	Energy efficiency programmes / activities
	DOE	Improve / increase energy affordability, reduce energy waste	General public, state officials/ PUC, utilities/ programme administrators, implementers, manufacturers (appliance standards)	Limited statutory authority outside of appliance standards, develop consensus in absence of mandate but risk going too far impairing function/use value, don't do behavioural research, just technology research	Public campaigns (Energy Hog), ENERGY STAR, educational information, work with programme administrators to see what they need to be more successful, authority, highlight where people are being successful (cache, prestige of authority), calculator tool on money spent to use appliance
	Regulator	Protecting ratepayer funds, deliver on savings goals that governor sets	Ratepayers, commissioners, intervenors	Not perceived as beholden / biased against the people they regulate (impartiality), competing priorities on budget/time	Influence, open to alternative approaches
Providers	Utility Planning	Provide forecasting activities, look at short term/long term time horizons, load research for budgeting and resource planning	Finance department, directors of grid planning, directors of customer programmes, regulators	Time management, cost and priority allocation, data challenges, management buy in of specific activities, reliable forecasts	Increasing staff, more internal resources for projects and provide services to the customer areas within utilities, improve analytics tools, increased budget for outsourcing, integrated resource planning
	Utility Programs	Educate customers about "always on" vampire energy behaviours and generate savings from increased awareness	Regulators, ratepayers, manufacturers, installers, marketing, experts	Funding mechanism to facilitate the programme (i.e. ratepayer dollars), customer energy literacy, cost of savings	Research to justify "always on" behaviour as wasteful and pose an opportunity for savings, programme design/management, trusted source of information for customers

	Manufacturers	Profitably, secure market share, secure more channels for products, secure floor space, price point set, follow corporate procedures (different mandates around new technologies and pricing)	Management, Board of Directors, retailers, distributors, shareholders, federal government (DOE/EPA), state agencies, utilities, customers, online channels, retail associations	Licensing bodies where applicable (DOE, CEE), distribution network, immigration (too many programmes in the same space), technology, research and development, money, market for it, how long would it take to develop the market, profitability, R&D schedule, mixed realities, interested customer base, products competition	ENERGY STAR, increased association involvement, marketing intervention/promotion, retailer/sales associate trainings, money, collaboration with utilities, more efficient distribution of products, building in technology that has energy efficiency modes already enabled, collaborate with a synergistic manufacturer that offers a complementary product/offering, brand cache, lobby
Middle Actors	Marketers	Effectively communicate objectives while maintaining customer satisfaction, consistent messaging, promote the company positively, brand awareness	Programme leads, customers, executive leadership, implementers, finance, experts on communication methods	Corporation communication schedule, keeping customer sales high/satisfaction high, keep within company branding guide, over communicating, time management	Brand awareness, forms of communication (mail, email, TV, radio, outreach events), partner with others, research on effective communication methods, marketing metrics
	Contractors/ Installers	Stay solvent/grow business, speed on the job while also performing well, installation-based on utility guidelines, increased sales, maintain contract	Utilities, utility customers, employees	Find quality installers (customer facing people) by performing background checks and paying more to get equipment installed, scheduling, household politics/issues education customers on how the technology should be used), forecasting from utilities, front end utility information	Trust and image with front facing employees
	Retailer	Generate increased profit, market share	Business, manufacturers, consumers	Shelf space, store point of view (hardware store vs. technology store), product variation, storage space both physical and technology for data, accounting, point of sales incentives	Trainer/educator for consumer on site, carry only the right products with energy efficiency potential, promote the energy efficiency products, choice architecture, online search viability, sales history
Experts	Academics	Provide unbiased/rigorous expertise to publish and communicate findings in a meaningful way, research done that produces papers and influences programmes	Peers (academic, implementation), funders/clients with expectations, subjects (do no harm)	Being publishable and having meaningful questions, overheads, intellectual property, access to quality data, time and funding source, make sure results are better than random, validity and controls, freedom to experiment vs. results that have already been accepted, funding source	Data-driven evidence (science, theory, empiricism), statistics, level of trustworthiness, impartiality, appropriateness

	EM&V Consultants	Producing credible evaluation research, making money, developing portfolio to market	Utilities/programmes (employers), regulators (read reports), company management	Staying in budget/on time, producing replicable research, programme design /implementation (data I have vs data I wish I had), project timeline (getting involved too late)	Research design practices/scientific methods, knowledge of other similar efforts evaluated, data collection tools, analysis/statistical techniques, peer network of other experts
Conscience	CEE	Help members identify the best practices, support CEE member needs, support replicability of programmes if clear	CEE members, DOE/EPA, CEE Board of Directors, CEE staff	Time, resources, balancing differences in members regulatory environments and what their regulators will find acceptable	Access to programme administrators, access to Task 24 resources, existing CEE resources (CEE *Behavior Program Summary*, etc.), respect as an independent / prestigious organisation
	Consumer Advocates	Represent/protect customer interests and obtain the lowest possible rates for consistent and safe services	Customers, regulators, researchers (sources of evidence)	Californian utilities are under the jurisdiction of the CPUC, understanding programme impacts/lifecycle vs. costs, systems thinking vs. cost effectiveness, staffing restrictions	Influence (cost, incompleteness), fear, influencing utilities and regulators

Table 2. End User Contexts.

BEHAVIOUR	**Goal: To reduce vampire energy waste in US & Canadian households**
TECHNOLOGY	• Less Technology/Appliances • Smart Power Strips, Smart Technology • Energy Efficiency Products, Sensors, Feedback Displays (Web or Widgets)
SOCIAL ASPECTS	• Children • Family Politics • Pets • Time • Peer Pressure (e.g. friends, colleagues) • Landlord • Safety • Upbringing/Tradition • Morals and Values • Social Media and (Fake) News
INFRASTRUCTURE	• Buildings • Ease of Access • Age and Equipment “Smarts” • AMI Meter Data and Interval Data • Information Communication Technology/ IoT, Standards/Regulations/Communication Protocols • Information Availability and Access • Interoperability/Supply Chain Availability
WIDER ENVIRONMENT (geographic, political, regulatory etc.)	• Socio-economic and Neoliberal Political System • Political Culture/ Governance • Separation Between Federal and State Governments • Standards and Codes • Paris Accord, Coalition of Mayors, etc. • Culture of Consumption • Community Culture and Social Norms • Geography and Climate • Education • Cost and Availability of Energy Efficiency Equipment

Issues and Insights

Following the initial *Behaviour Changer Framework* ("magic carpet") exercise and collecting stakeholder feedback at Workshop 1 in April 2018, the following challenges and insights were raised by participants:

- Claimable versus deemable savings via different programmes (behavioural, energy efficiency, demand response)
- Quantification and tracking of savings - are different methodologies accepted?
- Spill-over effect from schools to homes
- Smart home energy management as an interest area
- Time-of-use and dynamic rates, and the question of whether they are behavioural programmes
- Non-energy benefits, including the connection between health and energy usage, and whether they can be measured / included?
- Feasibility of scaling behavioural programmes beyond initial pilots
- Determination of the right time for targeted adoption
- Hourly data analysis to determine adoption rates, predicting highest peaks, targeting / segmentation, better baselines for actual usage rates, etc.
- Replicability and reproducibility in terms of data and persistence of savings.

2) US Top Issue: Evaluating Behavioural Programmes - Methodology, Credibility, Persistence & the Hard-to-Reach

Background

It became clear from the interviews with CEE sponsors (see below), that definitions of what constitutes behavioural savings programmes (as well as hard-to-reach customers) and thus, their implementation and evaluation, differed significantly between US States and Canadian Provinces. These differences were largely due to regulatory oversight and whether a utility was publicly- versus investor-owned. Some utilities had very close and trusted relationships with their regulators and were thus able to perform and claim for more varied behaviour change programmes and pilots. Others, notably California (see discussion from 2017 Karlin et al Draft White Paper below), had much more strict regulatory frameworks that applied to their large, investor-owned utilities (IOUs).

From Karlin et al (2017): Initial definitions of behaviour-based programmes in California limited claimable savings to a subset of behavioural interventions that employed: 1) comparative energy usage, 2) RCTs, and 3) ex-post measurement. Thus, in practice, behavioural savings programmes in North America (US and Canada) have been primarily focused on HERs that employ comparative feedback - this is particularly true for investor-owned (as opposed to publicly-owned) utilities. However, utilising a broader set of social science techniques in more innovative behavioural programmes could help tap into the estimated 20 percent savings potential from residential behavioural energy efficiency programmes. This White Paper identifies several opportunities to leverage more diverse behavioural strategies in order to enhance energy efficiency savings:

1. **Clarify the range of target behaviours available for intervention**
   It is important to keep the definition of energy behaviour sufficiently broad so as to capture the variety of behaviours that can impact energy savings. The IEA DSM programme's Task 24 states that "it is important not to confuse influences on behaviour, means of behaviour change, or the purpose of behaviour change, with what it is that is being changed..." (Mourik and Rotmann, 2013 p.10).

2. **Increase the types of strategies available for testing**
   While the current definition of behaviour in many jurisdictions limits programme administrators to comparative energy usage information - in other words, the use of social norms - the California IOU Behaviour Straw Proposal developed in 2013 expanded the definition from one technique (social norms) to instead mention seven distinct possibilities, additionally allowing for "other approved methods" even beyond those seven approaches. Energy efficiency programmes would benefit from perhaps even a further broadening of the scope of strategies to be used as this dynamic research field uncovers additional untapped behavioural strategies.

3. **Use a range of reliable research methods to infer causality**
   While RCTs are recognised as the gold standard for energy efficiency and behavioural savings programme evaluation, certain programmes are not practically evaluable using an RCT; thus, requiring RCTs limits the scope of behavioural programmes that are possible. Quasi-experimental methods have long been accepted by other disciplines as a reliable measurement approach in cases in which a RCT is either not possible or logistically unfeasible. These non-RCT techniques necessitate caution to prevent double-counting and require care to disentangle technology-related savings from behaviour savings but can be rigorous and practical tools when used appropriately.

Main Issues regarding evaluation of behaviour change programmes

Social marketing programmes previously selected behaviours to change by ranking them based on relative impact, probability and penetration, and then selecting the highest-ranked behaviours (McKenzie-Mohr, 2011). However, one CEE member utility selected behaviours by estimating actual energy savings over a twenty-year period of behaviours. The final programme behaviours were selected after compiling their residents' previous energy usage combined with a survey of likeliness of taking specific actions. Then, a global literature review was conducted to find the average uptake rates of those behaviours over twenty-years (Kassirer et al, 2014).

CEE *Behavior Program Summary*

The annual CEE Behavior Program Summary compiles CEE member programmes that include behavioural approaches throughout the US and Canada. It serves as a way to share information on programme details, programme components, evaluations, savings, and learnings. It should be noted, however, that this resource is fuelled entirely through voluntary data collection and is neither a census nor a random sample. Since this data collection shifted from one intended for the CEE membership only to a public report in 2012, 279 programmes have been reported from 94 organisations representing 41 US states and three Canadian provinces. A majority of the programmes reported have been for the residential sector (76 percent), with the commercial and industrial sectors representing 24 percent and 11 percent, respectively.

The CEE *Behavior Program Summary* includes the behaviour insights found in Figure 5. The majority of interventions are based on feedback, social norms, modelling behaviour and prompts. Using trusted messengers, interpersonal communication, priming and goal setting are also common intervention strategies.

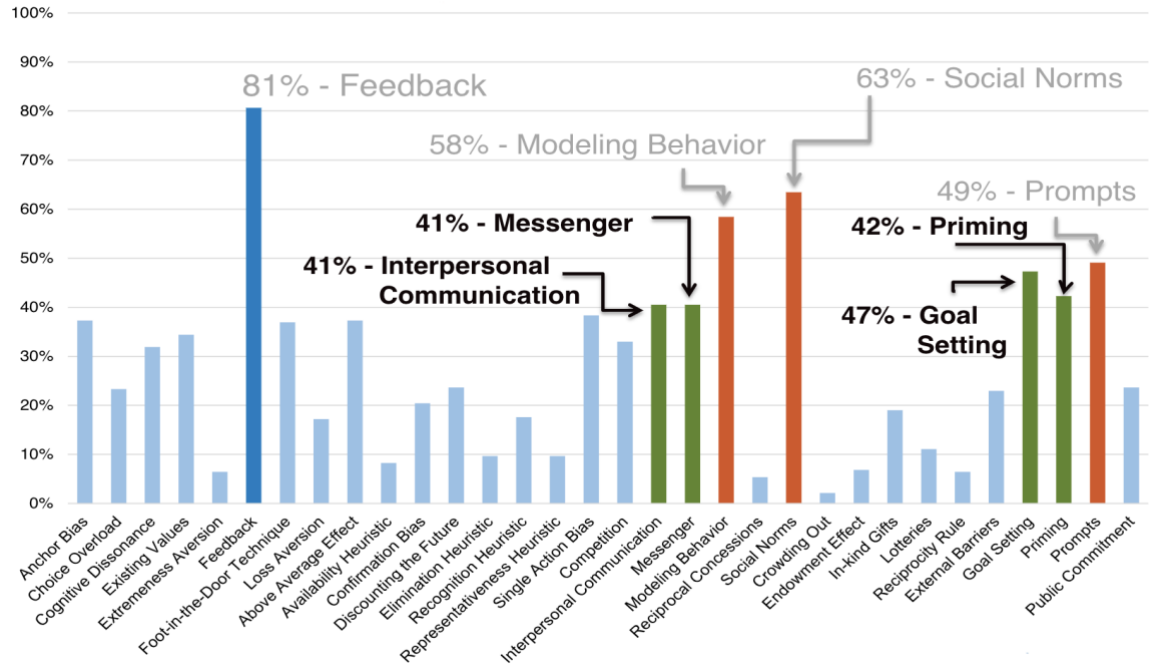

Figure 5. Most Reported Behavioural Insights, CEE Behavior Program Summary

The CEE *Behavior Program Summary* also collected data on evaluation, given that evaluation is an important aspect of programme implementation and can significantly impact whether or not a utility can claim savings and use behaviour-based approaches. Of the 279 programmes collected since 2012, 44 percent had completed the programme evaluation at the time of reporting, while 22 percent had an evaluation planned or underway. For the programmes that had either completed evaluations or had evaluations planned, a majority of those programmes received impact evaluations (97 percent) with just over half receiving process evaluations (60 percent). However, if only looking at programmes that were able to claim savings, there was a significant decrease in the number of overall programmes (see Fig 6).

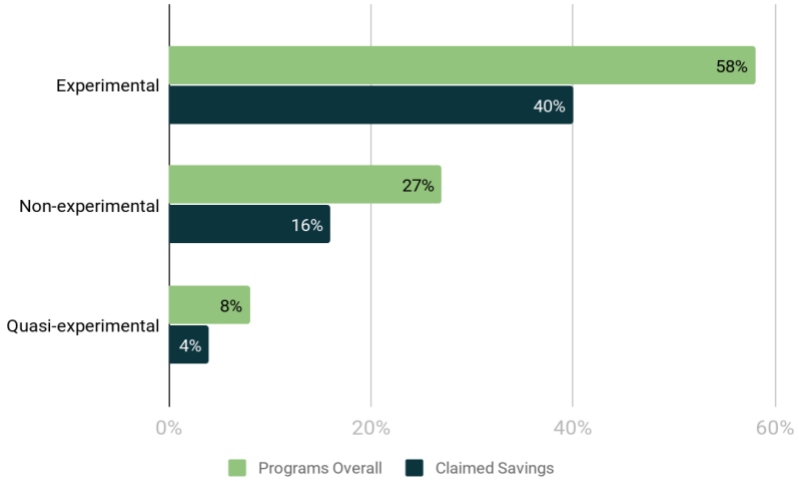

Figure 6. Evaluation Designs of CEE Member Programmes

The evaluation metrics commonly reported in the CEE *Behavior Program Summary* included: energy savings, customer satisfaction, awareness, cost savings, and number of participants reached. Energy savings was not the only metric for success, and was not used by all evaluated programmes, as illustrated in Figure 7.

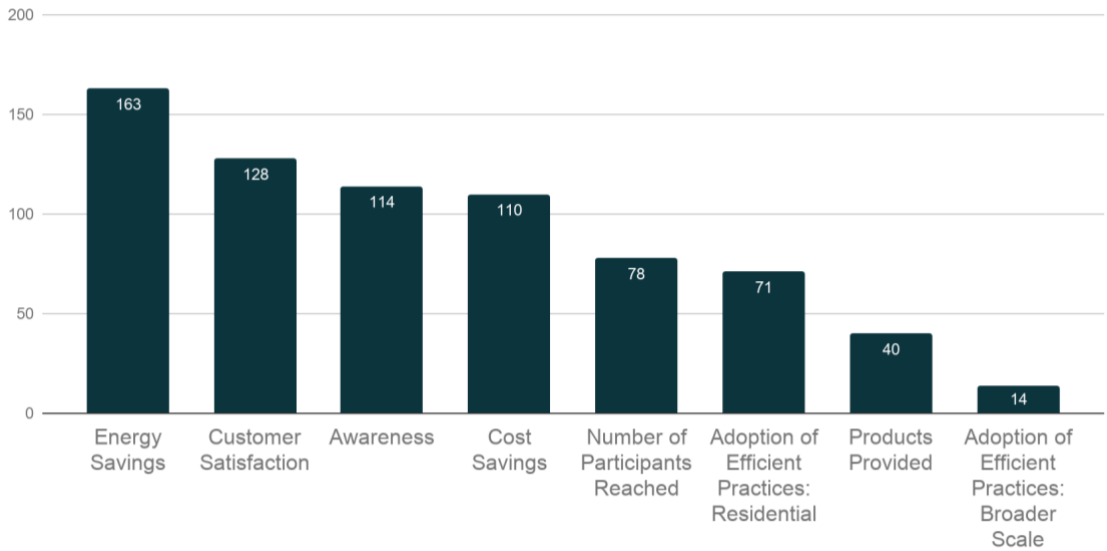

Figure 7. Evaluation Indicators of CEE Member Programmes

The CEE *Behavior Program Summary* has collected data on persistence since 2012. Of the 36 programmes that measured persistence, the average time the programmes assumed savings persisted was two years. It is important to note that this persistence information was provided across programme types, and it is reasonable to assume that persistence may vary widely based on the specific programme in question. However, given that persistence for behavioural programmes in the US and Canada is often assumed to be as little as one year, even this observational data suggests that a closer examination of behavioural persistence may support longer persistence assumptions for certain programme types.

A supplemental CEE members' report and public IEPEC paper on post-intervention persistence from HERs programmes found early indications that the longer HERs were distributed before the utility stopped sending reports to the given customer group, the longer the savings persisted. On average, across all HERs programmes examined for this analysis, there was a 78 percent persistence rate, and savings fully decayed four and half years after a programme's conclusion (Ashby et. al., 2017). That said, persistence rates may be far different for behavioural programme approaches outside the HERs model.

CEE sponsor input on defining and evaluating behavioural programmes

The semi-structured interviews with CEE sponsors allowed us to delve into the differences among sponsors in terms of the definitions, evaluations, and issues of behavioural programmes (See Appendix 2 for extended quotes from sponsors).

Defining a behavioural programme

Many sponsors noted that defining a behavioural programme depends on the underlying definition of behaviour itself. Sponsor A made the issue clear: "*It all depends on how you define "behaviour" [...] You would say everything we do is behaviour, and we'd say hardly anything is [according to] our regulator's definition.*" Indeed, some organisations consider behavioural programmes as "*those which require no capital investment [… and] are mostly focused on habitual behaviour change* (Sponsor H)," while other organisations "*increasingly acknowledge, in terms of programme design and budgeting, that behavioural programmes are broader* (Sponsor F)."

Despite varying definitions, on average, organisations offer one to two behavioural programmes and can usually claim savings for all their programmes. The most noted programme was HERs with six organisations (although two were discontinued due to low energy savings), followed by three strategic energy management (SEM) programmes, three programmes that could claim widgets and one kit programme. Other programmes were mentioned (such as school programmes or energy saving kits), but their savings could not be claimed.

Credible evaluation methodologies for behavioural programmes

A majority of our sponsors (six) use RCTs for claiming savings. One sponsor uses deemed savings, two use customer and site modelling, and another two incorporate randomised encouragement designs. Sponsor G said, "*programmes are designed with measurability and evaluation perspectives in mind*," which was reflected in the detailed responses most sponsors provided. Sponsors utilise a mix of in-house and third-party evaluators, and several expressed a desire to use behavioural and qualitative insights more extensively in order to undertake and scale up more ambitious pilots.

Energy efficiency or demand response programmes considered "behavioural" under the Task 24 definition

It became clear from discussion about the criteria and evaluation methodologies for behavioural programmes (above), that most utilities undertake many other interventions, which would be considered "behavioural" under the Task 24 definition but may not be in their own jurisdictions. In the US and Canada, these interventions are often undertaken by different teams internally, and have different evaluation methodologies and strategies from programmes meeting the internal definition of behaviour.

When asked to describe interventions that are not considered behavioural internally, but fit the Task 24 definition of behavioural, sponsors focused on measurement methodologies. For one organisation, "*behaviour is around measured changes in energy usage whereas other [energy efficiency or demand response] programmes are counting widgets (*Sponsor F)." Evaluations are commonly conducted by third party evaluators, and deemed values are the most common savings measure and are sourced from white papers published by the regional energy commissions. Multiple sponsors acknowledged limitations of a strict binary delineation between behaviour and non-behaviour savings programmes - as one sponsor asked, "*Where does behaviour start and end?*" Other sponsors mentioned efforts to include behaviour in analysis techniques of savings programmes: "*We are looking at hybridising methods to support the quantitative data on widget numbers with behavioural insight s[...]*" (Sponsor F).

Regulatory challenges and possible solutions

The degree to which regulation affected utilities varied widely. Some experienced somewhat stricter requirements, definitions and frameworks, while others (typically not IOUs) had very few, if any, challenges related to regulatory oversight. Common regulatory challenges included:

- Regulators' level of knowledge about behavioural approaches (mentioned twice)
- Stakeholder perceptions (twice)
- Proving and earning acceptance of new methods and their cost-effectiveness (four times)
- Budget constraints (twice)
- Very specific or narrow definitions of behaviour (three times)

One sponsor (I) stated the problem simply as: "*Proving that we are not going to waste ratepayer money is hard*." Multiple sponsors expressed interest in advancing and expanding creative and novel savings programmes: "*It's certainly possible to establish longer term goals, especially in terms of climate change, health and social benefits, and also a more holistic view towards energy efficiency." (*Sponsor C).

Sponsors also discussed multiple promising opportunities to meet regulatory requirements. Sponsor A said, "*I think a movement first toward adoption of national evaluation standards within the US will go a long way in overcoming these challenges*." Four sponsors advocated for starting small scale pilots to show proof of concept and then scale up. Other opportunities included:

- Educational meetings with the Regulator (mentioned twice)
- Building trusted relationships with the Regulator (three times)
- Using insights gained from international programmes (once)
- Including non-energy benefits (NEBs) (twice).

Persistence of behavioural savings programmes

Generally, many sponsors had not measured persistence yet because their programmes were too new, but nearly all sponsors expressed interest in the topic or planned to measure persistence in upcoming years. Three organisations noted they were currently measuring persistence and have shared their reports with CEE. CEE published a summary of the CEE *Behavior Program Summary* findings on persistence in 2017 (Ashby et al., 2017); although the final report is not public, there is an IEPEC paper that summarises the persistence findings that is public.

Hard-to-Reach customers

CEE sponsors' definition of and experience engaging with hard-to-reach customers

Hard-to-reach generally refers to targeted groups of customers "*who are difficult to recruit or design programmes for"* (Sponsor G). However, some sponsors consider the term itself inaccurate because communicating with, and engaging customers are different goals: "*We send a utility bill to everyone, so presumably there is at least one communication channel in place, but creating engagement and action is more difficult"* (Sponsor H). Moreover, internal definitions of hard-to-reach customers vary widely. Some definitions include underserved communities (including rural, non-white, or low-to-moderate income customers), anyone who doesn't participate in programmes, small and medium businesses (SMBs), or multifamily buildings. Several sponsors also noted the multi-faceted challenge of successfully engaging customers in rental units; depending on the situation, the person who pays the bill and makes decisions regarding energy efficiency can be the landlord, owner, tenant, or a combination thereof.

Regulator's definition of hard-to-reach customers

The California Public Utility Commission (CPUC) has struggled with the challenge of defining hard-to-reach for decades. One challenge to defining these groups is they are not static; for example, in the late 1990s and early 2000s, residential and small commercial customers were underserved relative to large businesses, which benefited disproportionately from utility energy efficiency programmes. However, as of 2013, the CPUC defined hard-to-reach residential customers as "*those customers who do not have easy access to programme information or generally do not participate in energy efficiency programmes due to a language, income, housing type, geographic, or home ownership (split incentives) barrier."* In addition, hard-to-reach *markets* include multifamily, low to moderate income, and small businesses.

Interview with a former CPUC Regulator

As a consequence of the difference in ownership (investor vs federal) between the US and Canada and the international community, the regulatory environments are similarly disparate. North American regulators are charged with ensuring ratepayer funds are used responsibly and efficiently; therefore, energy efficient programmes must be evaluable and effective. To better understand and accurately depict the American regulatory environment, specifically that of California, the Task 24 Operating Agent undertook a semi-structured, hour-long interview with a former CPUC Regulator. This individual has worked in the field of evaluating energy efficiency programmes for almost fifteen years, and she identified a number of timely topics from the regulatory perspective.

Risk Reward Incentive Mechanism

In 2005, the CPUC implemented an experimental programme called the Risk Reward Incentive Mechanism (RRIM). The goal was to motivate utilities to improve and expand energy efficiency programmes via a system of incentives and penalties. Ultimately, from the perspective of this Regulator, the RRIM encountered several obstacles that limited its effectiveness and fostered distrust between utilities and the CPUC.

The former Regulator was involved in the initial development of the RRIM and explained it was a tiered incentive structure such that utilities who met or exceeded an 80 percent savings goal received a cash incentive, those who saved 60 - 80 percent received nothing, and those whose savings fell below 60 percent received a penalty. The penalty component of the RRIM was criticised when it was implemented since, at the time, penalties were not typically recommended in the US for failing to meet energy efficiency targets.

Efficiency programmes were evaluated by the CPUC post-implementation, and the evaluation results were not available to stakeholders (such as implementers and utilities) until three years after deployment. There were large discrepancies between the savings utilities had claimed and the evaluation results, triggering an adjustment of any penalties or rewards received. However, the variables driving the discrepancies were often not within utility control. A report titled "Raising the stakes for energy efficiency: California's RRIM" highlights why this created disputes: "*Because of the sharp earnings cut-offs, large differences in incentive pay-outs to utilities did not necessarily reflect meaningful differences in performance [...]"* In one such case "*Pacific Gas & Electric would have either been awarded $180 million or penalised $75 million for the energy efficiency efforts already finished, depending on whose interpretation of programme evaluations you used.*"

Pay for Performance and NMEC programmes

The former Regulator wrote a lengthy blog about performance-based efficiency, which drew on her experiences during the contentious RRIM, and also addresses the newer Pay for Performance and Normalised Metering Energy Consumption (NMEC) programmes. She cautions all players to learn from history:

> "*All actors in the system need to know how they will be measured and judged to assume responsibility and maintain accountability, and they need the ability to monitor progress as things roll out. Methods need to be consistent, but also transparent and replicable. To be replicable, the nitty-gritty execution of the method must be accessible. If there is too much room for interpretation, methods will be vulnerable to persistent attack, usually from the party that didn't get a favourable outcome. Energy efficiency requires upfront agreement, because there is no other ground truth.*"

She warns that the energy-efficiency savings calculations that are "endlessly disputed" makes scaling the system increasingly difficult and impedes efforts to "create and exchange a commonly understood unit of savings." When using the agreed-upon evaluation methods, "*the method and execution of that method must be accessible at the design phase, foundational to the deployment effort, and serve as the contractual basis for payment*."

The former regular noted several criteria necessary to achieve high savings:

- Creativity
- Clear rules (i.e. clarity about the bar against which the programme/portfolio is being measured)

- Support for utilities to incorporate creative IoT / behaviour approaches into programmes
- Support for grid and integrated-resource planning funded by investors advocating for climate change and green energy issues, including insurance companies
- Transparency and trust
- An understanding of the risk from investing in given programmes, and options to minimise that risk.

Summary of Issues and Insights from CEE Sponsors and Regulator Interviews

Several main regulatory issues and insights emerged from considering the perspectives of both regulators and programme administrators:

- There is a spectrum of definitions of a "behavioural savings programme"; the Task 24 definition is most broad and the initial CPUC definition is the most limited. Most sponsors' programmes lie somewhere in between.
- Most sponsors undertake other energy-savings programmes which do have behavioural aspects (e.g., efficiency programmes based on deemed widget values or demand response programmes applying time-of-use rates).
- Best practices for measuring and evaluating behavioural savings programmes are not clear. In some states, RCTs are the primary or only acceptable evaluation methodology. In others, sponsors only need to prove "cost-effectiveness", and can use a variety of allowable methods.
- The use of third-party evaluators and implementers is very common, though in-house EM&V expertise is desirable and considered very cost-effective.
- A close, trusted relationship with the Regulator is also highly desirable and can lead to greater flexibility for behavioural pilot interventions and claims.
- Intra-organisational barriers to wider adoption of behavioural measures exist. These include pressure to prove the value of qualitative or behavioural measures, time, budget or resource constraints, and a need to avoid "reinventing the wheel."
- Proving persistence of energy savings is crucial and highly desired, but currently lacking; many programme administrators are interested, but most have not had the chance to test the persistence of their programmes. The 2017 CEE report on persistence is still the main standard.
- Definitions of hard-to-reach customers are highly variable, both within and between the residential and commercial sectors. There is a terminology issue with the terms "hard-to-reach" versus "underserved." As several sponsors pointed out, customers aren't hard-to-reach as long as you can send them their utility bill (and they pay it). Underserved may be a more holistic description of the many user groups who do not engage with current behavioural or energy efficiency programmes. The IEA DSM Executive Committee (ExCo) approved a draft concept paper for a new project on these "hard-to-reach" customers. The project will begin in 2019, pending final IEA DSM ExCo approval in April 2019.
- Pay for Performance programmes "*can aid in assessing savings and motivating persistence in savings from complex, multi-measure efficiency projects including those with behavioural or operational changes, where it is difficult to estimate savings in advance based on average historical data.*"[2]
- California's history with shareholder incentives and energy efficiency programme evaluations offers a number of lessons:
    - Trust is imperative and enables collaboration between all the actors in the system (utilities and their customers, regulators, implementers, and evaluators).
    - Shareholder incentives are useful but must be implemented carefully within the complex utility operational environment.
    - Earnings calculations should not be too sensitive to small changes in energy savings estimates.
    - High-stakes incentives such as the RRIM merit caution and should include institutional arrangements for energy savings measurements and dispute resolution processes that are accepted by all parties.
    - Although conflict of interest concerns prevents utilities from leading the evaluation process that determines incentive earnings, a transparent view of the process and utility participation in development of evaluation methods is still warranted.
    - The increasing availability of real-time energy usage data is a promising opportunity for behavioural programme evaluation.

[2] https://www.nrdc.org/experts/merrian-borgeson/can-paying-performance-increase-energy-savings

3) International Comparison

Literature review

A short literature review was undertaken in July 2018 by the Task 24 Operating Agent, in Elsevier's *Scopus* (the largest abstract and citation database of peer-reviewed literature). Keywords used were behaviour change, evaluation, evaluation methodology, persistence, hard-to-reach, and regulatory barriers. The reference lists for any papers that appeared relevant were checked for further, relevant references (a so-called "backwards search"). Thirty-nine papers of the total 176 papers were included as relevant to the CEE sponsors.

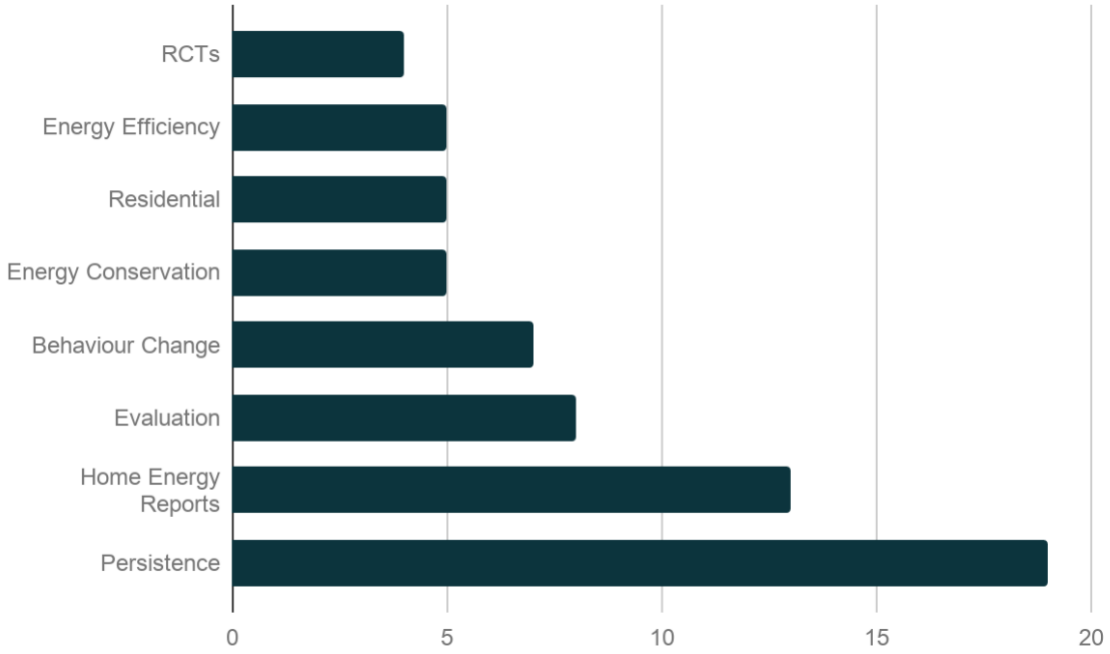

Figure 8. Top Keywords from the Literature Review.

Of the 39 papers deemed relevant to CEE sponsors, the most common keyword was persistence, as illustrated in Figure 8. In addition, twenty-one of the papers measured persistence, fourteen provided evaluation methodologies, three addressed hard-to-reach customers, and two identified regulatory barriers. A searchable database of these papers was provided to the CEE sponsors.

Country Comparison: The EU Energy Efficiency Directive

The Energy Efficiency Directive (EED) is the main policy instrument at the EU level to reach the 20 percent energy savings goal by 2020. Article 7 is a key pillar of the EED and requires Member States (MSs) to introduce energy efficiency obligation schemes (EEOSs). Under the EEOS, energy companies must save an annual 1.5 percent of their energy sales via additional energy efficiency projects. A paper by Bertoldi et al. (2015) outlines the "*EEOS [are] introduced and planned by MSs in terms of sectoral coverage, obligated actors, eligible projects, monitoring and verification (M&V), baseline and additionality, sanctions, trading rules if any, and public authorities' role. A comparison among the different national EEOS is made, and their common features highlighted. Key issues including the time scale needed to introduce an effective EEOS, type and number of obligated partners, changing business models of energy companies and scale of expenditure are discussed.*"

In terms of supplier obligations, Bertoldi et al. (2015) detailed that in the 1990s, energy companies started to provide energy efficiency solutions and financing through DSM programmes. In the 2000s, the role of the energy companies became regulated by law and targets were introduced. The regulations and targets are a "*market-based policy oriented towards end-use energy efficiency based on energy-savings quotas (obligations) for some categories of energy market operators (usually energy distributors or suppliers).*" The energy savings are either verified by the regulator (or national authority) or the savings

are certified via "white certificates" (certificates for energy savings). These white certificates in national schemes are used as part of an energy efficiency trading system, and if the parties are not subject to energy-saving quotas, they can still certify the eligible projects but then sell the white certificates to generate additional revenue streams while also increasing the white certificate market liquidity.

In theory, this trading system can minimise the overall costs of compliance for the obligated parties because their "*savings targets are set relative to their number of customers/amount or value of energy sold, and the dominant market players have the greatest obligations."* For example, in Denmark, the new energy efficiency market has enabled the electricity distribution companies to establish energy service companies demonstrating innovation under the EEOSs. The Danish example illustrates how EEOSs have to fit into the "*existing policy landscape and are often used to meet multiple goals, rather than simply delivering energy savings."*

In addition, EEOSs can be used to:

- Transform the business model of energy companies
- Develop the supply chain and improve the installation quality of particular measures
- Prioritise vulnerable customers or encourage new actors to enter the energy efficiency market.

Non-compliance with the EU EEDs can lead to quite significant consequences. For instance, the Dutch government was sued, by over 100 Energy Performance Advisor companies and the EU Commission, due to failure to properly implement Article 7 of the EU's Building Performance Directive. Article 7 addresses the "*right of an (interested) buyer or renter of a building to be informed about the energy performance of the home or building by means of Energy Performance Certificate, which includes an objective, professional advice on the costs, benefits and effects of measures that will improve the energy performance."* The goal of the Directive is to promote building energy savings because a crucial moment to encourage savings is when a home gets a new owner or user.

For more information on EU utility obligations, see also:

Bertoldi, P., Rezessy, S., Lees, E., Baudry, P., Jeandel, A., and Labanca, N. (2010). Energy supplier obligations and white certificate schemes: Comparative analysis of experiences in the European Union. *Energy Policy* 38 (3), 1455–1469.

EU, 2006. Directive 2006/32/EC of The European Parliament and of the Council of 5 April 2006 on energy end use efficiency and energy services and repealing Council Directive 93/76/EEC.

EU, 2012. Directive 2012/27/EU of The European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/ EU and repealing Directives 2004/8/EC and 2006/32/EC.

Rosenow, J., Platt, R., Flanagan, B. (2013). Fuel poverty and energy efficiency obligations – A critical assessment of the supplier obligation in the UK. *Energy Policy* 62, 1194–1203.

Task 24 Case Study Analysis

In the spirit of international comparative studies, Task 24 focused on a case study analysis specifically in its Phase 1 Subtask 1 (*Helicopter overview of behavioural models, theories, and disciplines*) and Subtask 2 (*In-depth Case Studies from Participating Countries*). For this report, we summarised all Subtask 1, 2 and 6 (*Top DSM Issues of each Participating Country*) case studies into a case study summary database, highlighting main aspects of each case study. Links to reports and other relevant information were also captured, including details on whether persistence and hard-to-reach customers were studied. In addition, we found as many primary reports as were still available for each of these case studies, which are provided in a separate database. The most relevant reports were summarised in more detail in another database, addressing the specific topics chosen by the CEE sponsors. During Workshop 2 (see below), four Task 24 case studies were used to highlight how to follow good social science and design thinking processes.

The following Task 24 case studies and reports may be of particular interest given the CEE sponsors' selected topic (see summary descriptions in the detailed summary database):

General descriptions of case studies, evaluation methodologies and metrics

1. Mourik and Rotmann (2013). Most of the time what we do is what we do most of the time. And sometimes we do something new. Subtask 1
2. Karlin et al. (2015). What Do We Know About What We Know? Subtask 3, Deliverable 3.
3. Mourik et al. (2015). Did you behave as we designed you to? Subtask 3, Deliverable 3A: Positioning paper on monitoring and evaluation.
4. Van Summeren et al. (2015). From 'I think I know' to 'I understand what you did and why you did it'. Subtask 3 Deliverable 3B.
5. Batey and Mourik (2016). From calculated to real energy savings performance evaluation: how ICT-innovation and do-It-yourself user-generated data monitoring can enable to improve real performance evaluation of energy efficiency initiatives, *Energy Efficiency* 9(4), 939–950.
6. Karlin et al. (2015). Exploring Deep Savings: A Toolkit for Assessing Behavior-Based Energy Interventions, *IEPEC Conference,* Long Beach, USA.
7. SCE (2015). Subtask 9 - Dimensions of Energy Behavior: Psychometric Testing of Scales for Evaluating Behavioral Interventions in Demand Side Management Programmes.

Evaluation methodologies prescribed by a Regulator

8. Grazer Energieagentur (2017). Ergebnisbericht. [in German]
9. Kallsperger and Rotmann (2017). Subtask 6&7 - Final Report Austria.

Different evaluation methodologies for specific case studies

10. Lang (2014). Subtask 2 - The Energy Hunt Austria.
11. Goodhew et al. (2015). Making Heat Visible: Promoting Energy Conservation Behaviors through thermal imaging. *Environment and Behaviour* 47(10).
12. Rettie et al. (2013). CHARM Project. 'Is social norms marketing effective? A case study in domestic electricity consumption', *European Journal of Marketing*.
13. Mukai et al. (2016). Evaluating a Behavioral Demand Response Trial in Japan: Evidence from the Summer of 2013. *Energy Efficiency* 9(4), 911–924.
14. Batey and Garcia (2014). "Innovative Use of ICT to uncover space-usage patterns & improve energy-efficiency." *EESAP5 Proceedings*.
15. Rotmann and Chapman (2018). Subtask 9 - Using Bayesian Modeling to Evaluate Home Energy Saving Kits in Ireland.
16. Rotmann (2018). Evaluating Auckland Council's HEAT Kit.
17. SEAI (2018). Subtask 6&7 - Final Report Ireland.
18. Cowan et al. (2017). CHS case study: Designing a successful behaviour change programme for hospital building staff.
19. Cowan et al. (2018). It's Not my Job: Changing Behavior and Culture in a Healthcare Setting to Save Energy. *ACEEE Summer Study,* Monterey, US.

Non-energy benefits and hard-to-reach energy users

20. Various evaluation reports and IEA (2014). The Warm Up New Zealand: Heat Smart Programme.

Persistence of behaviour change interventions

21. Löbsiger-Kagi et al. (2018). Does it last? Long-term impacts of an app-based behavior change intervention on household electricity savings in Switzerland. ERSS 47:16-27.

Task 24 Expert Interviews

Short, written interviews were also conducted with three of the national experts of Task 24 (Austria, Sweden and New Zealand). They all operate under quite different regulatory regimes and their answers reflect their views about their country's contexts. These interviews can be used to further contrast how behavioural interventions and programmes are distinct in different countries. (See Appendix 3 for extended quotes from the three National Experts.)

From Teresa Kallsperger, Austrian Task 24 Expert

Energy utilities, under the Austrian Energy Efficiency Act (EEffG), are mandated to save 0.6 percent of their previous year's energy sales via energy efficiency interventions. EEffG's goal is to *reduce the energy end use through the energy efficiency interventions in the utility sector by 159 Petajoules (PJs) by 2020*; *of the 0.6 percent of the utilities' annual energy savings, 40 percent needs to be contributed by Austrian households.*

Some of the legislated residential interventions are outlined (in German) here, which shows the breakdown of utility interventions: 34 percent were in heating and hot water interventions; 25 percent in retrofitting; 7 percent in lighting; 6 percent in mobility interventions. However, there is uncertainty around which of the interventions actually led to energy savings; for example, 30 percent of individual interventions were not clearly defined in the monitoring document but claimed to have led to 44 percent of the total energy savings (Adensam et al., 2013; see also Kallsperger and Rotmann, 2017; and Task 24 Endbericht, 2017 [in German, summary in the literature database]). In addition, Austrian utilities commented that in some cases it is cheaper for them to pay the penalties for not achieving their targets than it is to implement energy efficiency interventions.

Mehmet Bulut, Swedish Task 24 Expert

"Behaviour change is regarded as very important to achieve the Swedish climate goals and the government has therefore introduced several incentives", for example free hourly metering is offered. In addition, there is a growing interest in nudging and inter-disciplinary research that incorporates the social sciences. In Sweden, behavioural savings programmes encourage energy efficiency behaviour either through the uptake of energy-efficient products or through influencing usage patterns; these behavioural programmes are similar to other energy efficiency or DSM programmes.

To evaluate these behaviour programmes, evaluators use a combination of quantitative and qualitative surveys (i.e. data measurements and surveys). Regulation does have an impact on the behavioural savings programmes, but they need more ways to keep end users engaged; they have an ongoing cooperation with the Energy Market Inspectorate ("Inspectorate") to help address programme barriers. These barriers are mostly related to policies from policymaking institutions, such as the Ministry.

The Swedish Energy Agency has a strong interest in seeing savings persist, so they will be participating in the next Task on hard-to-reach energy users. However, Sweden, unlike other countries, has a very low energy poverty rate, but there are issues related to senior citizens living alone in old, large, energy inefficient houses and securing electricity supply in some rural areas. These rural areas are therefore considered hard-to-reach because they are very dependent on electricity. To make these areas less dependent, the Swedish Energy Agency works to promote a completely renewable, sustainable energy system with a highly secure supply.

New Zealand's Utility Obligations

The main government organisation in charge of energy efficiency and conservation efforts in New Zealand is the Energy Efficiency and Conservation Authority (EECA). It releases the *New Zealand Energy Efficiency and Conservation Strategy* (NZEECS) every five to six years, in addition to the Ministry of Business, Innovation and Employment's *New Zealand Energy Strategy* (NZES). The targets under NZEECS are:

1. Decrease in industrial emissions intensity (kg CO_2-e/$ Real GDP) of at least 1 percent per annum on average between 2017 and 2022.
2. Electric vehicles make up 2 percent of the vehicle fleet by the end of 2021.
3. 90 percent of electricity will be generated from renewable sources by 2025 (in an average hydrological year), providing security of supply is maintained.

None of the targets are binding or monitored using third party evaluators.

The electricity regulator is the Electricity Authority (EA). New Zealand is regarded as having one of the most deregulated electricity markets (see Ministry of Business, Innovation, and Employment (MBIE), 2015). Utilities are part of the New Zealand Emissions Trading Scheme but have no direct energy efficiency obligations. Only 0.5 percent of a customer's utility bill is spent on "Market Governance" which includes costs for energy efficiency programmes and to run the EA (see EA, 2018). The EA runs a scheme called "What's my Number" to educate consumers about the benefit of comparing and switching

retailers. EECA runs programmes and information campaigns to reduce energy use in the residential, industrial, commercial and transport sector. Main campaigns are the *Warm Up New Zealand* insulation programme (Mourik and Rotmann, 2013) and various Minimum Energy Performance Standards and Labelling schemes. EECA also has a "Rightware Tool" programme to help consumers compare energy efficiency of different appliances. New retail pricing plans include time-of-use tariffs that let consumers and tools monitor their own consumption online. Customers can then increase savings by shifting their electricity use to off-peak times or take advantage of lower night rates to charge their electric vehicles.

Dr. Daniel Gnoth, PowerCo, New Zealand National Expert

The New Zealand current energy portfolio programme does not take much of a priority because the EECA strategy focuses more on information about purchase behaviours. In addition, recently, "*behaviour-focused programmes also haven't been awarded to any of the large research funding bids.*"

New Zealand defines energy-savings programmes that leverage behaviour change as *all initiatives across the value chain, with utilities as 'Middle Actors' who rarely engage with end users directly*. In New Zealand, utilities may interact indirectly with end users through web-based applications that leverage pricing signals and information provided by the utility sector. In terms of other energy efficiency or DSM programmes, the initiatives differ based on pricing or policy which provide signals and incentives to the market; in some cases, the other programmes are information-based.

To evaluate these behaviour programmes, utilities measure energy savings or load shifting in kW, but they still require more granular information about the customers because customer satisfaction is still very important, especially in terms of employing additional initiatives and maintaining engagement. Regulation provides a *framework by which industry can engage with their customers*. Typically, there are not many barriers to the programmes, aside from policy design, because utilities work closely with regulators. In addition, persistence is not monitored for most programmes in New Zealand, and utilities often depend on interventions that do not engage consumers directly.

Task 24 international workshops and expert feedback

Task 24 finalised Phase II with two international expert workshops, at the two largest behaviour change conferences of 2018. It ran a four-hour workshop at the BEHAVE conference in Switzerland on September 5, 2018, and a 90-minute special session at the BECC Conference in Washington D.C. on October 10, 2018. In both, the main topics chosen by CEE sponsors were addressed (see Appendix 4 for the summarised feedback from each workshop).

BEHAVE conference, Zürich Switzerland, September 5-7, 2018

In Switzerland, we used audience prompts and prepared handouts, 30 of which were returned by the almost 60 attendees, who hailed from 13 different countries. Summary findings of these responses show that the majority of respondents (23/30) came from the research sector.

BECC conference, Washington D.C., October 7-10, 2018

At the BECC conference session, after a short overview of seven years of Task 24 findings, and the CEE involvement in Year 7, we convened a CEE sponsor panel who answered audience questions, and prepared questions around the topics of how many of their behavioural programmes could claim savings, how to evaluate them, and related regulatory barriers and solutions. The main problems discussed were **restrictions on what behavioural savings programmes** were able to be undertaken and claimed and **how they were evaluated**; and if there were any **regulatory barriers to implementing behaviour change programmes**.

Task 24 Workshop 2 - Washington D.C., October 10-11, 2018

Following the Task 24 / CEE special session at the BECC conference, we convened for a one and half day workshop with the CEE sponsors at the US DOE headquarters in Washington, DC. On Day 1, the Task 24 Operating Agent provided an overview of the work and results undertaken to date and sponsors discussed in-depth current work and future topics of interest (including the new IEA DSM Hard-to-Reach energy users Task).

On Day 2, our Project Partner Dr. Beth Karlin from SCI, led us through a workshop on the *SCI Process for Behaviour Change* (see Figure 9 below), including worksheets to be filled in during the day. Detailed Workshop minutes are available to CEE sponsors.

Figure 9. Different phases of the SCI process for behaviour change.

We used the following Task 24 case studies to highlight the different phases of the SCI process and where improvements could be made using better social science and design thinking processes:

Green Leasing in Commercial Office Buildings in Sweden (see Rotmann and Bulut, 2018; BELOK 2017 and 2018; Janda et al., 2017).

Atrium Health Building Operators in Hospital Buildings in the Southern United States (see Cowan et al., 2017 and 2018).

Warm Up New Zealand: Heat Smart residential insulation subsidy programme (see Mourik and Rotmann, 2013; IEA, 2014; MOTU, 2014).

Irish Home Energy Saving Kit Programme using different Middle Actors (see SEAI, 2018; Rotmann and Chapman, 2018; Rotmann, 2018c and 2018d).

Issues and Insights

- There is a lot of primary literature on evaluation, but as our 2015 Methodology Review by Karlin et al. has shown, not much of it clearly outlines evaluation methodologies, nor do many papers give the actual survey questions or scales that were used for qualitative analysis. The result is that programme administrators often need to reinvent the wheel in evaluating their behavioural programmes.
- In addition, not much primary literature is relevant to utilities directly, as the work undertaken is usually more academic and smaller in scale than most North American utility behaviour change programmes.
- Studies on persistence and hard-to-reach energy users are harder to find in the primary literature than more general behaviour change intervention studies.

- In terms of CO_2 emissions, behaviour of individuals in the US has been found to be single-highest emissions sector, yet has received much lower research funding relative to other energy efficiency topics such as energy-efficient technologies (Vandenbergh et al., 2010). This is also reflected in the amount spent by utilities on behavioural savings programme funding versus broader energy efficiency or DR programme funding.
- The EU, US, and Canada all have regulations that require the utility sector to implement energy efficiency and behaviour change interventions and prove energy savings via prescribed methodologies. However, there are some specific differences, particularly around consequences of non-compliance (in the US it is more incentive-based versus in the EU it is more penalty-based) and motivations (in the US behavioural interventions are often based on financial motivations, and in the EU the appeals are often more on moral grounds). A country like New Zealand, with a highly-deregulated utility sector, has no comparative regulations, only government-set strategies and targets with little consequences for non-compliance.
- Given the large variability across the US in terms of the regulatory frameworks and barriers and what constitutes a behavioural savings program, any cross-country comparison thus has to be taken with a grain of salt. In looking from North America (US and Canada) to other countries, it is important to consider and understand the different underlying contextual factors. This report can only scratch the surface of such factors.
- There are significant differences in research approaches to behaviour change in the EU and New Zealand versus North America (US and Canada), particularly around the emphasis on individual versus societal behavioural changes. This includes different disciplinary approaches and models e.g. a preponderance of psychological theories such as the *Theory of Planned Behaviour* in the US versus sociological theories such as *Practice Theory* in the EU and also more social-marketing focused approaches like *Community-Based Social Marketing (CBSM)* or neoclassical economic models like the *Deficit Model* in the US. However, in both the EU and US, behavioural economics and behavioural insights have found a strong niche and support by policymakers and industry alike. Many of the most commonly used approaches to design and evaluate behavioural interventions are not regarded as the most useful by socio-technical researchers (see Sovacool, 2017).
- Even though European experts seem to have a stronger mandate to undertake "purely" behaviour change research and pilots, they still mirror some of the issues around credibility the CEE sponsors have raised. Some concern lack of funding and support compared with energy efficiency technology research and programmes; some are around engineers and policymakers being dubious about the validity of qualitative data and evaluation and lack of inclusion of social scientists at the design stage. At the BEHAVE conference, we also had an interesting discussion around social justice and class issues, particularly when addressing the hard-to-reach or fuel poor sector (this will be expanded upon in the new HTR Task).
- A lot of these issues can be overcome by following proper social science and design thinking processes, which can be internationally-validated and become something of a standard of how to undertake and evaluate behaviour change interventions. Our Project Partner, SCI, workshopped such a process with our CEE sponsors, using Task 24 case studies to illustrate rigorous practice in real-life interventions.

4) Main questions remaining

From Sovacool (2014): "*Research questions that arise include:*

1. *How can the benefits of "human-centered" research methods be best coupled with quantitative forms of data collection and analysis?*
2. *Human-centered, as well as "mixed" research methods, tend to be more complex (difficult to fit into a box), expensive, and subjective than others—how can they be improved?*
3. *How can researchers minimise bias—their own, and that of their subjects—when doing research?*
4. *How does one balance depth—going "deep" on a small sample of respondents—with breadth—producing conclusions sufficiently generalisable beyond those respondents?*
5. *Where does one draw the line between a research method, a way of collecting data, and a conceptual framework, a way of framing or filtering data?"*

Which Behaviour Changers could be better engaged to overcome some of the programme and evaluation barriers identified by CEE sponsors?

Given the unique US regulatory context, which insights from international behavioural programmes may be the most relevant?

Who are the actual End Users for improved behavioural programme evaluation? Is it Utilities? Utility customers? Regulators? Policymakers? Third Party Evaluators or Implementers?

What are the best evaluation methodologies for behavioural programmes, here and overseas? How can credibility of an intervention be more effectively established?

What are the best ways to measure persistence of a behavioural intervention? How can we ensure that the practice of measuring persistence in new programmes becomes more common?

Recommendations

The table below summarises the recommendations for the top DSM issue, which we focussed on in the second phase of Task 24 in North America (US and Canada). In general, to solve any behavioural intervention, all relevant Behaviour Changers should align, collaborate, and communicate with each other and with the End Users whose behaviours are being targeted for change. As several CEE sponsors have highlighted, educating and fostering trusted relationships with their regulators means that they have much greater flexibility in terms of what behaviour change programmes they can implement and claim savings for, and how they can prove cost-effectiveness. Continuous learning from international best practice and cutting-edge approaches to behaviour change evaluation is recommended to all experts in this space, as is sharing of data and insights, including those approaches that have been less effective.

Table 3: Recommendations for Behaviour Changers.

Behaviour Changers	Recommendations to increase approval of behavioural interventions by better and more credible evaluation methods
Decision-makers (e.g. Regulators) 	- Try to be inclusive of non-RCT evaluation methods for behavioural programmes when savings for a given programme type are well established or alternate methods are more appropriate for a given programme approach. - Consider opportunities to co-create with utilities' new programmes that include behavioural techniques to allow buy-in from both parties upfront. - Support evaluation of positive side effects of DSM interventions, such as NEBs; develop guidance around how NEBs can be evaluated and valued.
Experts (e.g. 3rd party evaluators) 	- Develop further evaluation methods for relevant DSM interventions, especially non kWh-ones (NEBs). - Provide non-RCT evaluation recommendations for regulators based on proof-of-concept from case studies and pilot projects. - Continue to publish best practice examples and case studies, including in shared, non-academic literature. - Develop standard, validated process of how to 'do' behaviour change in practice and how to evaluate those efforts.
Providers (e.g. Utilities) 	- Learn about proven behaviour change and evaluation methods outside of HERs and RCTs. - Consider shifting language from "behaviour programmes" to "behavioural tools" and or "behavioural processes". - Triangulate quantitative and qualitative data to improve behavioural insights - Develop evaluation methods for relevant DSM interventions, especially non kWh-ones (NEBs). - Provide recommendations for regulators and decision-makers from proof-of-concept from case studies and pilot projects and consider opportunities to co-create with regulators new programmes that include behavioural approaches. - Share learnings and data, collaborate with other Providers and Experts to create standard methods, processes and databases (e.g. energy saving tips library), and look to peers abroad for new approaches suited for domestic testing. - Support and include trusted Middle Actors (e.g. contractors) and Conscience (e.g. environmental or community groups) to help approach energy end users, especially ones that are "hard-to-reach". - Measure persistence of behaviour change programmes in terms of energy savings but also measure qualitative insights into how energy users responded
Middle Actors (e.g. CEE, contractors) 	- Support end users and Providers when implementing DSM interventions (with technology and consulting via being the trusted intermediary). - Help evaluate the positive side effects of DSM and energy efficiency. - Engage in training programmes for behaviour change, educate beyond technology-focused solutions.
Conscience (e.g. NDRC) 	- Strengthen evidence-based communication of the topics climate change and energy efficiency, and providing education to End Users in a way that best practice research has demonstrated is most likely to be effective. - Engage and collaborate with other *Behaviour Changers* to promote proof-of-concept, and verbalise the positive social and environmental impacts from better behaviour change interventions, including NEBs.

Conclusions

One undercurrent throughout US participation in Phase II of Task 24 is that there are cultural and societal differences between the US and the European countries from which many of the Task 24 case studies were drawn. As a result, approaches that were successful in Europe cannot simply be transplanted to the US. That said, the absence of a strict regulatory structure abroad that determines which types of behavioural approaches may and may not be explored means that European peers have been more easily able to test out how innovative behavioural approaches can be used to shift energy usage behaviours. Although these approaches would not necessarily be transferable to the US without modification, they provide a glimpse into what behavioural techniques may be appropriate for adaptation or testing in the US. With this caveat in mind, there were a number of key findings from this work:

Behavioural Terminology

Some utilities in the US have shifted away from referring to behavioural programmes as such in order to avoid concerns about the durability of achieved behaviour changes. If the goal of behavioural programmes is to move beyond deemed measures such as the purchase and installation of hard measures, then there needs to be some distinction between acquiring new energy-efficient equipment and taking actions that reduce energy usage. One approach that could help address this challenge is a concerted effort to shift the language used to describe these programmes across the industry; rather than referring to "behaviour programmes," instead referring to the *behavioural tools and processes* that can effectively be used to enhance all programmes across sectors.

Persistence

Given that human behaviour is more variable than energy-efficient equipment, there's a greater need to understand and measure persistence for behavioural programmes - or, in other words, what happens to a programme's achieved behaviour changes and related energy savings after the programme activities cease. It is noteworthy that there has been more research in the US relative to other countries on the topic of persistence. Most research on persistence in the US has examined how rapidly energy savings decline after energy reports are no longer sent to customers participating in Home Energy Report (HER) programmes. Initial research from eight utilities suggests decay rates of energy savings ranged from 2 percent to 30 percent per year, with many programmes reporting annual energy savings decay in the 20 percent range during the first year (Ashby, K., Gutierrez, V., Menges, S., Perich-Anderson, J. (2017): Keep the Change: Behavioral Persistence in Energy Efficiency Programs). These findings demonstrate that behavioural persistence is tangible and measurable. However, further research is needed to understand whether these persistence findings apply in other programme contexts outside of HERs.

Programme Scale and Programme Origins

Throughout this work, a few substantial differences emerged between energy efficiency programmes in the US and Europe. One key difference between programmes across the continents that emerged early on in this effort was that programmes in the US are typically conducted on a much larger scale relative to their European counterparts; in the US, programme participation is often on the order of tens of thousands of customers as opposed to hundreds of customers. As a result of the smaller number of participants typically involved in European energy efficiency efforts, it is more challenging to determine causality for resulting energy savings in a way that would be perceived as credible in the US. European countries do not often evaluate their programmes in the same way as is typical in the US, because it is often not required. Additionally, in Europe, energy efficiency efforts are often spearheaded by federal government entities, whereas in the US, energy efficiency programme administrators - typically, but not always - implement programmes. As a result, European programmes often weight non-energy benefits (NEBs) more heavily than their peers in the US, by simple virtue of who is tasked with implementing and measuring the value of the programme.

Behavioural Programme Evaluation

On the evaluation side, there was a disparity between which programmes were considered behavioural programmes in the US, and which programmes were being evaluated as behavioural programmes (e.g. RCTs). Currently, the RCT is most commonly used to evaluate behavioural programmes in the US. When qualitative evaluation methods are used, it is typically for *process evaluations* as opposed to *impact*

evaluations. There may be great value in expanding evaluation approaches for behavioural programmes to include other techniques, such as qualitative methods, quasi-experimental designs, Pay for Performance, etc.

Looking ahead, further gains in improving behaviour change efforts in the US can be achieved by continuing to develop enhanced non-RCT behavioural programme evaluation methods, shifting the rhetoric around behaviour change away from "behavioural programmes" and towards "behavioural tools and processes", and by continuing to measure the persistence of successful behaviour change efforts.

References

Ashby, K., Gutierrez, V., Menges, S., & Perich-Anderson, J. (2017). Keep the Change: Behavioral Persistence in Energy Efficiency Programs. Paper presented at International Energy Program Evaluation Conference, Baltimore, US.

Ashby, K. CEE Behavior Program Summary (2018).

Batey, M. and J.P. Garcia (2014). "Innovative Use of ICT to uncover space-usage patterns & improve energy-efficiency." EESAP5 Proceedings.

Batey, M. and R.M. Mourik (2016). From calculated to real energy savings performance evaluation: how ICT-innovation and do-It-yourself user-generated data monitoring can enable to improve real performance evaluation of energy efficiency initiatives, Energy Efficiency 9(4), 939–950.

BELOK (2018). Subtask 6&7 – Collaboration and Green Leasing: A case study of the Swedish Energy Agency's new office building in Eskilstuna. IEA DSM Task 24.

Bergold, J. & Thomas, S. (2012). Participatory Research Methods: A Methodological Approach in Motion. Forum Qualitative Social Research, 13(1).

Bertoldi, P., Rezessy, S., Lees, E., Baudry, P., Jeandel, A., and Labanca, N. (2010). Energy supplier obligations and white certificate schemes: Comparative analysis of experiences in the European Union. Energy Policy 38 (3), 1455–1469.

Bertoldi, P., Castellazzi, L., Oikonomou, V. and I. Moorkens (2015). How is Article 7 of the Energy Efficiency Directive being implemented? An analysis of national energy efficiency obligation schemes. ECEEE Summer Study 2015, Hyeres, France.

California Public Utilities Commission (CPUC). ENERGY EFFICIENCY POLICY MANUAL Version 5 July 2013.

California Public Utilities Commission (CPUC). Decision (D.) 12-05-015 – minimum definition of behavior programs.

Cobben, D. (2017). Subtask 6 & 7: Case Studies NL Higher Education and ICT. IEA DSM Task 24.

Cowan, K., Sussman, R., Rotmann, S., & Cox, R. (2017). Subtask 11: Real-life case studies Designing a Behaviour Change Programme for Hospital Facilities Staff. IEA DSM Task 24.

Cowan, K., Sussman, R., Rotmann, S. and E. Mazzi (2018). It's Not my Job: Changing Behavior and Culture in a Healthcare Setting to Save Energy. ACEEE Summer Study Monterey, US.

Churchouse, T., Mahoney, L., & Rotmann, S. (2015). Closing the Loop – Behaviour Change in Demand Side Management. Paper presented at UKERC The Meeting Place, London.

Dietz, T., Gardner, G.T., Gilligan, J., Stern, P.C., and M. P. Vandenbergh (2009). Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. PNAS 106 (44): 18452-18456.

EU, 2006. Directive 2006/32/EC of The European Parliament and of the Council of 5 April 2006 on energy end use efficiency and energy services and repealing Council Directive 93/76/EEC.

EU, 2012. Directive 2012/27/EU of The European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/ EU and repealing Directives 2004/8/EC and 2006/32/EC.

Goodhew et al. (2015). Making Heat Visible: Promoting Energy Conservation Behaviors through thermal imaging. Environment and Behaviour 47(10).

Grazer Energieagentur (2017). Ergebnisbericht. [in German]

Hanleybrown, F., Kania, J., & Kramer, M. (2012). Channeling Change: Making Collective Impact Work. Stanford Innovation Review.

Hantrais, L. (1995). Comparative research methods. Social Research Update, 13 (Summer), pp. 2-11.

International Energy Agency (2014). Capturing the Multiple Benefits of Energy Efficiency. Paris, France: IEA.

IEA DSM Task 24 Phase I: Closing the Loop: Behaviour Change in DSM – From Theory to Practice.

IEA DSM Task 24 Phase II: Behaviour Change in DSM – Helping the Behaviour Changers.

IEA DSM Task 24 (2013). Subtask 1 - Definitions.

Janda, K., Rotmann, S., Bulut, M. and S. Lennander (2017), Advances in green leases and green leasing: Evidence from Sweden, Australia, and the UK, ECEEE Summer Study Proceedings, Hyéres, France.

Kania, J. & Kramer, M. (2011). Collective Impact. Stanford Social Innovation Review, Winter.

Kallsperger, T. and S. Rotmann (2017). Subtask 6&7 - Final Report Austria. IEA DSM Task 24.

Kallsperger, T. & Rotmann, S. (2018). Policy Brief Austria. Graz, Austria: IEA DSM Task 24.

Karlin, B., Ford, R., Wu, A., Nasser, V., & Frantz, C. (2015). Subtask 3 Deliverable 3A: What do we know about what we know? A Review of Behaviour-Based Energy Efficiency Data Collection Methodology. Los Angeles, USA.

Karlin, B., Ford, R., & McPherson-Frantz, C. (2015). Exploring Deep Savings: A Toolkit for Assessing Behaviour-Based Energy Interventions. Paper presented at International Energy Programme Evaluation Conference, Long Beach, USA.

Karlin, B., Ford, R., & McPherson-Frantz, C. (2016). Evaluating Energy Culture: Identifying and validating measures for behaviour-based energy interventions. Paper presented at International Energy Programme Evaluation Conference, Amsterdam, Netherlands.

Karlin, B., Lupkin, L., Ford, R., Bunten, A., Forster, H., Zaval, L. (2017). California Behavioral Definition: Review and Recommendations.

Kassirer, J., Korteland, A., Pedersen, M. (2014). "Team Power Smart Sparks Increase in Low-Priority, Repetitive Behaviors." Social Marketing Quarterly 20(165).

Lang, G. (2014). Subtask 2 - The Energy Hunt Austria. IEA DSM Task 24.

Löbsiger-Kagi et al. (2018). Does it last? Long-term impacts of an app-based behavior change intervention on household electricity savings in Switzerland. ERSS 47:16-27.

McKenzie-Mohr, D. (2011). Fostering sustainable behavior (3rd ed.). Gabriola Island, Canada: New Society.

Ministry of Business, Innovation and Employment (MBIE) (2015). Chronology of New Zealand Electricity Reform.

Measham, T.G. (2009). Social learning through evaluation: a case study of overcoming constraints for management of dryland salinity. Environmental Management 43(6), pp. 1096-1107.

Monitoringstelle: Rechtliche Grundlagen für Energieeffizienz in Österreich. Retrieved from https://www.monitoringstelle.at/

Moezzi, M., Janda, K., & Rotmann, S. (2017). Using stories, narratives, and storytelling in energy and climate change research. Energy Research and Social Science, 31, pp. 1-10.

Moore, L., de Silva-Sanigorski, A., & Moore, N. (2013). A socio-ecological perspective on behavioural interventions to influence food choice in schools: alternative, complementary or synergistic? Public Health Nutr. 16(6), 1000–1005.

Mourik, R.M. & Rotmann, S., (2013). Subtask 1: Most of the Time what we do is what we do most of the time. And sometimes we do something new. Analysis of case studies. Eindhoven, NL and Wellington, New Zealand: IEA DSM.

Mourik, R.M., van Summeren, L.F.M., Rotmann., S., & Breukers, S. (2015). Did you behave as we designed you to? Monitoring and evaluating behavioural change in demand-side management: from what to why. Paper presented at ECEEE Summer Study, Hyéres, France: ECEEE.

Mukai et al. (2016). Evaluating a Behavioral Demand Response Trial in Japan: Evidence from the Summer of 2013. Energy Efficiency 9(4), 911–924.

Österreichische Energieagentur (2013). Methoden zur richtlinienkonformen Bewertung der Zielerreichung gemäß Energieeffizienz- und Energiedienstleistungsrichtlinie 2006/32/EG: Bottom Up Methoden.

Rettie et al. (2013). CHARM Project. 'Is social norms marketing effective? A case study in domestic electricity consumption', European Journal of Marketing.

Rosenow, J., Platt, R., Flanagan, B. (2013). Fuel poverty and energy efficiency obligations – A critical assessment of the supplier obligation in the UK. Energy Policy 62, 1194–1203.

Rotmann, S. and R.M. Mourik (2013). Closing the loop between theory, policy and practice: IEA DSM Task 24 on behaviour change. ECEEE Summer Study, Hyeres, France.

Rotmann, S. & Mourik, R.M. (2014). Work Plan for Phase II: IEA DSM Task 24 Behaviour Change in DSM – Helping the Behaviour Changers. Wellington, New Zealand: IEA DSM.

Rotmann, S., Goodchild, B., & Mourik, R.M. (2015). Once upon a time… telling a good energy efficiency story that 'sticks.' Paper presented at ECEEE Summer Study, Hyéres, France: ECEEE.

Rotmann, S. (2016). How to Create a 'Magic Carpet' for Behaviour Change. Paper presented at ACEEE Summer Study on Energy Efficiency in Buildings, Monterey Bay, US: ACEEE.

Rotmann, S. (2017). Once upon a time…Eliciting energy and behaviour change stories using a fairy tale story spine. Energy Research and Social Science, 31, pp. 303-310.

Rotmann, S. (2018a). Subtask 8 – Toolbox for Behaviour Changers: Behaviour Change from A-Z. IEA DSM Task 24.

Rotmann, S. (2018b). Evaluating Auckland Council's HEAT Kit. IEA DSM Task 24.

Rotmann, S. (2018c). Subtask 6&7 – Cross-Country Case Study Comparison Ireland – Home Energy Saving Kit Library Programmes. IEA DSM Task 24.

Rotmann, S. (2018d). Subtask 6&7 – Final Report New Zealand. IEA DSM Task 24.

Rotmann, S. and D. Chapman (2018). Subtask 9 - Using Bayesian Modeling to Evaluate Home Energy Saving Kits in Ireland. IEA DSM Task 24.

SEAI (2018). Subtask 6&7 - Final Report Ireland. IEA DSM Task 24.

Sovacool, B. K. (2014). What are we doing here? Analyzing fifteen years of energy scholarship and proposing a social science research agenda. Energy Research and Social Science, 1, pp. 1-29.

Sovacool, B.K. and D.J. Hess (2017). Ordering theories: Typologies and conceptual frameworks for sociotechnical change. ERSS 47 (5).

Southern California Edison (2015). Dimensions of Energy Behaviour: Psychometric Testing of Scales for Evaluating behavioural Interventions in Demand Side Management Programmes.

Van Summeren, L.F.M., Mourik, R.M., & Rotmann, S. (2015). Subtask 3 Deliverable 3B: From "I think I know" to "I understand what you did and why you did it." Eindhoven, NL: IEA DSM.

Vandenbergh, M.P., Stern, P.C., Gardner, G.T., Dietz, T., & Gilligan, J.M. (2010). Implementing the behavioral wedge: Designing and adopting effective carbon emissions reduction programs. Environmental Law Reporter, 40(6), pp. 10547-10554.

Walzer, N., Weaver, L., & McGuire, C. (2016). Collective impact approaches and community guideline developments. Journal of Community Development, 47(2), pp. 156-166.

Appendix 1. Task 24 Overview, Objectives and Deliverables

Background and Overview

Task 24 research is aimed at improving DSM and sustainable energy use by influencing human behaviour. During Phase I (2012 - 2015), the teams in the different participating countries focussed on translating behavioural theory into practice. They built a network of >250 behaviour change experts who made an inventory of available theories, models and approaches, gathering over 60 practical examples and case studies from 20+ countries (for more details, see Rotmann 2016a).

Main lessons learned from Phase I (see Mourik and Rotmann, 2013):

- There are a variety of applicable theories and models that are currently underutilised when designing behavioural interventions (especially from sociology and multidisciplinary studies);
- There is much to be gained by using combinations of approaches, and moving from the current, overwhelmingly technocratic approaches to consider more 'human' perspectives. This includes fostering and facilitating multi-stakeholder collaborations;
- Many of the collected stories and case studies showed a lack of in-depth understanding turning behavioural theory into practice and a clear need of further field research and validated tools;
- Most countries had not clearly prioritised their top behavioural DSM issues for further research, or failed to include all relevant stakeholders (*'Behaviour Changers'*) in the selection process;
- There were some top behavioural DSM issues in each country where the theory from Phase I could be turned into best practice in Phase II, using *Participatory Action Research* (PAR) approaches (e.g. see Bergold, 2012).

In 2015, Task 24 continued with a new Phase II based on these insights. First, the national teams selected their countries' top-priority areas in behaviour change in DSM (**Subtask (ST) 6** – "The Issues"). This selection of top areas was performed with the IEA DSM ExCo member of each participating country, the appointed *National Experts* and other country experts (*Behaviour Changers*). The DSM priorities differed between countries, as did their (technical, economic, political and societal) potentials and risks due to different national contexts. We will ascertain and highlight these country differences, using policy briefs in **ST 10** ("Overarching Story").

After having identified the top priority areas for energy efficiency within a country, one area was selected for further research in detail. Once the top areas were chosen in each country, the national teams brought the relevant *Behaviour Changers* together to explore the key issues supporting and hindering the uptake of DSM in the current system (**ST 7** – "The People"). The key systemic issues were then explored in facilitated multi-stakeholder workshops. Finally, in some countries, we could then engage the relevant *Behaviour Changers* in designing a "real-life intervention" (**ST 11**). We also developed more focused intervention approaches and a "Toolbox for Behaviour Change" (**ST 8**) as well as a "Beyond kWh" evaluation tool (**ST 9**). The latter are discussed in depth elsewhere but will be mentioned here in their application in **ST 6 & 7**.

The major hypothesis of the Task 24 Phase II approach is that *a Collective Impact Approach* (Kania and Kramer, 2011) *which fosters collaboration among a variety of stakeholders - together with whole-system visualisation exercises in participatory action research settings, and using storytelling as overarching 'language' - will lead to more successful behavioural interventions where multiple benefits to the end users and each* Behaviour Changer *can be clearly evaluated*.

Objective of Task 24

The main objective of Phase II is to take good theory into practice to allow *Behaviour Changers* to:

- Engage in an international expert network - **ST 5** 'THE EXPERTS'
- Identify the most appropriate DSM themes to focus on - **ST 6** 'THE ISSUES'
- Identify and engage countries' *Behaviour Changer* networks for at least one of the top three DSM themes - **ST 7** 'THE PEOPLE'
- Use and test a *CIA* to develop shared methodologies, guidelines and a common 'language' based on narratives to aid *Behaviour Changers* - **ST 8** 'THE TOOLS'
- Standardise how to evaluate behaviour change programmes 'Beyond kWh' and 'Beyond Energy' including multiple benefits analysis - **ST 9** 'THE MEASURE'

- Collate national learnings into an overarching (international) story to understand, compare and contrast the different behaviour change approaches, risks and opportunities and which recommendations can be universally applied - ST 10 'THE STORY'.

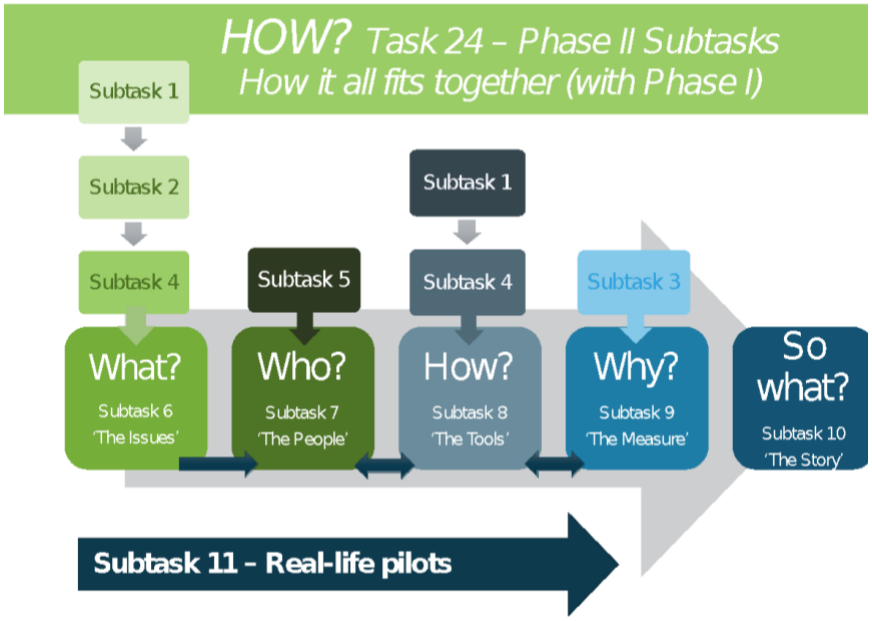

Figure 1. Overview of how Subtasks fit together (In Phases I & II).

Methodology of Task 24

We describe the individual approaches used in ST 6 & 7 in the US, in more detail below. The overarching tools that were developed and tested in Task 24 Phase II (ST 8 & 9) are summarised first.

Subtask 8 - Main tools in the Toolbox for *Behaviour Changers*

The toolbox (Rotmann, 2018a) has a strong focus on tools that support the appropriate context for the *Behaviour Changers* and are more conducive to developing systemic interventions, with stories and case studies illustrating their application. The workshop sessions with the *Behaviour Changers* focused on testing these tools on a variety of countries, sectors, contexts and behavioural issues.

Objectives

- Use the *CIA* to unite *Behaviour Changers* from five main areas on a specific DSM issue (both chosen in ST 6 & 7).
- Collect information for a decision-making tree to pick the most appropriate case studies and models of understanding analysed by Task 24 (ST 1, 2 & 6).
- Develop the common language of storytelling further and provide different examples of using storytelling and narratives in practice and how to best do it in the specific areas of focus and in each of the *Behaviour Changers'* sectors.
- Identify the tools in each *Behaviour Changer's* Toolbox of Interventions, analyse their pros and cons, risks and opportunities, where they fall short and how another tool from another *Behaviour Changer* could overcome this deficit.
- Continued testing and development of the Evaluation Tools (ST 3 & 9) that can prove if a (toolbox of) intervention/s leads to actual, ongoing behaviour changes in practice. The *Behaviour Changers* will give feedback on its potential applicability, risks and additional needs by working through (hypothetical or real life) examples chosen in ST 6 and using double-loop learning approaches to assess multiple benefits of interventions.
- Collaborative development of a testable Toolbox of Interventions for each top DSM focus area, where each *Behaviour Changer* sector has clearly identified and measurable roles and responsibilities. This intervention may then be taken into a real-life setting and trialled in practice (either as ST 11 or outside of Task 24).

- The toolbox is built on national and sectoral context specificities but will be synthesised and tested (e.g. in international conferences - ST 5) for the general aspects that are of international validity (ST 10 - Overarching Story).

Deliverables

D 12: Testable toolbox for *Behaviour Changers*. This includes:

- A description and evaluation of the validity and effectiveness of the *CIA* in the energy arena, as a peer-reviewed paper (Rotmann, 2016a and b).
- A decision-making tree that enables *Behaviour Changers* to better utilise the findings of ST 1 & 2 (de Zeuuw, 2018).
- A peer-reviewed paper on the impact of storytelling in energy research (Rotmann, 2017; Moezzi, Janda and Rotmann, 2017).
- A collection of sector stories from each *Behaviour Changer* sector (Rotmann, 2017).
- This includes different behavioural intervention tools each *Behaviour Changer* has at their disposal in each of their national and sectoral contexts (see US workshop minutes).
- Continued testing and development of evaluation tools created in ST 3 & 9 (Rotmann and Chapman, 2018a and b).
- Testable toolbox for national *Behaviour Changers* when choosing to take part in ST 11 (see Cowan et al. 2017 and 2018) and/or synthesis of internationally valid tools to feed into the Overarching Story (ST 10, to be published).

Storytelling

We discussed the importance of language, definitions and jargon, and need to clearly define it, above. We also needed to find an overarching 'language' in order to bridge the many different disciplines, sectors and *Behaviour Changers* we were dealing with: this language was *storytelling*.

The Task thus embarked on a journey of using various narratives and storytelling tools to simplify learnings, bridge silos and 'translate' between different *Behaviour Changers*. Some of the approaches are discussed in Rotmann, Goodchild and Mourik (2015). The main Task 24 approach of using a fairy tale story to elicit stories from 100s of *Behaviour Changers* in over 20 countries was detailed in a Special Issue on "Narratives and Storytelling in Energy and Climate Change Research" in *Energy Research and Social Science* (Rotmann, 2017). Task 24 Operating Agent Dr. Sea Rotmann co-edited this Special Issue with Drs. Mithra Moezzi and Kathryn Janda (see Moezzi, Janda & Rotmann, 2017 for an introduction and summary). 35 excellent papers are showcased in this Special Issue, which forms the ultimate collection on storytelling in energy and climate change research to date. The introduction to the Special Issue became the number one most downloaded article in the Energy Research & Social Science Journal in 2018.

The "Collective Impact Approach"

Task 24 uses two different, yet complementary, approaches to facilitate multi-stakeholder collaboration in the more practice-oriented Phase II: *The Collective Impact Approach* (Kania and Kramer, 2011) and the *Behaviour Changer Framework* (Rotmann, 2016a). The *CIA* was first developed to help social entrepreneurs deal with complex social problems. This approach, aimed at long-term social change, proposes a collective, rather than an individual approach for solving difficult problems. Walzer et al. (2016) argue that complex situations which would normally be difficult to solve, can be solved using the CIA. This CIA is described by Collaboration for Impact as: *"…an innovative and structured approach to making collaboration work across government, business, philanthropy, non-profit organisations and citizens to achieve significant and lasting social change*."

Five conditions are listed that are needed to create such a collective impact (Figure 2):

1. A common agenda,
2. Mutually-reinforcing activities,
3. A shared measurement system,
4. Continuous communication and
5. A backbone support organisation.

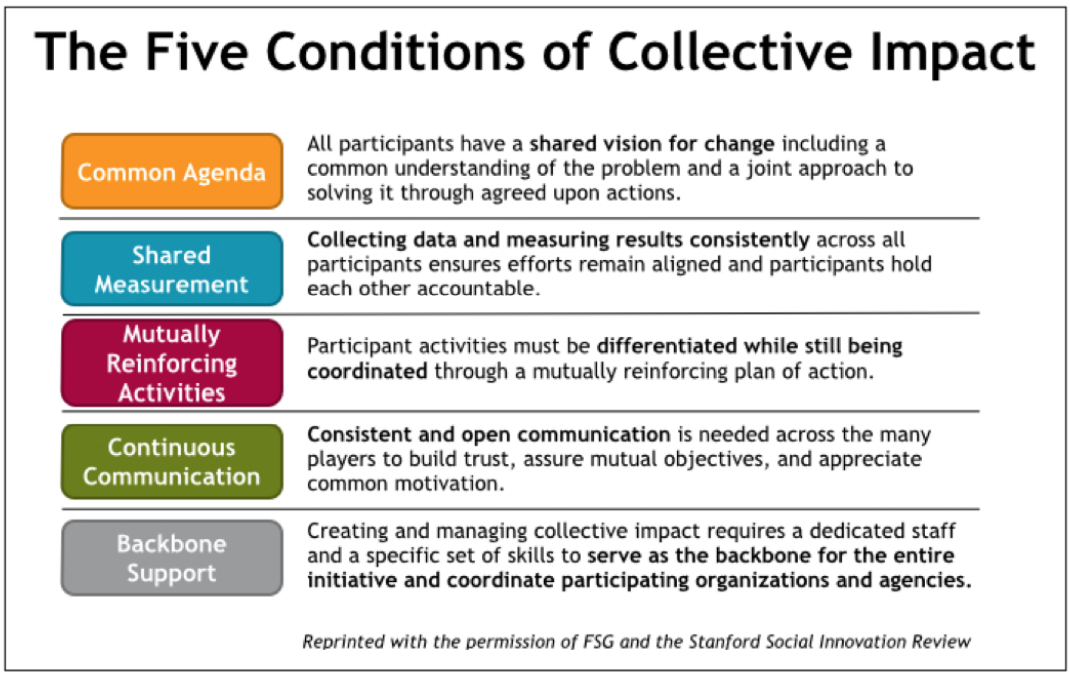

Figure 2. The five conditions of the CIA (from Kania and Kramer, 2011).

A **common agenda** is important to create a common understanding of the problem and the solution in order to make sure all *Behaviour Changers* agree on taking the same road to the common goal. Secondly, it is also important that the relevant *Behaviour Changers* perform **mutually-reinforcing activities**, making sure that they do not impede other *Behaviour Changers* or their stakeholders. Thirdly, it is also important that there is a **shared measurement system** so that outcomes of all *Behaviour Changer's* actions are measured and reported in the same way in order to share and learn from each other. Furthermore, to create trust and a common vocabulary, it is of high importance that actors **communicate continuously**. Lastly, a separate **backbone support organisation** needs to be created that facilitates a change of mind set, creates publicity and mobilises resources. Kania and Kramer (2011) explain that backbone organisations are especially important for providing direction, facilitation of the dialogue, mobilising funding and handling all the different layers of linked collaboration. *Behaviour Changers* are interdependent on each other, on other stakeholders, and they also operate in different and sometimes very complex contexts when confronted with various political, financial and social pressures. Their mandates may be insufficient to affect large-scale behaviour change, when in direct conflict to it. Hence, complex problems that include technical, organisational, social and behavioural dimensions ask for a way to collectively address challenges. In order to do so successfully and to enable shared learning, a trusted *Facilitator* and 'translator' is crucial (e.g. Measham, 2009). In Phase II, Task 24 took on these important roles.

The CIA offers a way to implement change via a top-down and bottom-up mixed approach. Most research on this approach focuses on situations in which a collective impact is created by organisations that are independent units. The first version of the CIA did mention the five principles on which successful collective impact should be based. However, nothing was said on further steps that should be taken or what institutions could function as backbone organisations. In 2012, the CIA wrote a second article in which they remedied both shortcomings. Hanleybrown, Kania and Kramer (2012) state that there are three phases that have to be fulfilled for creating collective impact.

- In the first stage, **action has to be initiated**. In order to do so, the landscape of the social problem has to be understood first and a **champion** has to stand up. The importance of champions is to take care of attracting financial resources and creating a sense of urgency, striving for collaboration.
- It is also important to **organise for impact**. This means that common goals, a shared measurement system and backbone organisation have to be arranged.
- In the third and last phase **action has to be sustained** and impact should arise. Active learning and coordination are described to be essential for success (ibid).

For more detail on how the *CIA* is utilised in Task 24 and how it can be assessed in real-life applications, see e.g. Cobben (2017) and Cowan et al. (2017).

The Task 24 *Behaviour Changer* Framework

To create a more hands-on tool to identify and work on the five conditions of the CIA, Task 24 developed the so-called "*Behaviour Changer* Framework", which was later dubbed "the magic carpet of behaviour change" by a major US utility during a Task 24 workshop. This framework was created to provide a visual overview of the social ecosystem, focusing on all relevant stakeholders, i.e. the *Behaviour Changers* from the different sectors and their relationships with one another, and the *End User*. This framework focuses on a chosen issue (ST 6) from the perspective of the *End Users* and their behaviour, as well as their context in terms of technology, social aspects, infrastructure and the wider environment (including political and regulatory). It also focuses on each of the *Behaviour Changers* in the system, what their main mandates, stakeholders, restrictions and tools are, and how they interact with one another and with the *End User* (for detailed description of the process and actor types, see Rotmann 2016a). We used this framework to play through the chosen issue of "vampire energy in the residential sector" (see US workshop minutes), at the first US Task 24 workshop in San Francisco, April 8, 2018.

An alternative view of our Energy System

An important point of departure from the current technocratic view of the Energy System is that in Task 24, we pose that *our energy system begins and ends with the human need for the services derived from energy (warmth, comfort, entertainment, mobility, hygiene, safety, etc.). We suggest that behavioural interventions using technology, market and business models, as well as changes to supply and delivery of energy are the all-important means to that end*.

The Behaviour Changer *Framework* operates on a different 'model of understanding' of the energy system, one based on *behavioural socio-ecology* (e.g. Moore, de Silva Sanigorski & Moore, 2013). The socio-ecological framework encourages both whole-system interventions, and also the explicit understanding of how more focused interventions might depend on factors at other levels (including the various human actors in a given system) for their effectiveness, acceptability or sustainability to be achieved (ibid, p1002). Here, this means:

1. Exploring the views, values and experiences of the various experts and decision-makers engaged in a given 'energy socio-ecosystem' (often also including the energy *End User* whose behaviour they are ultimately trying to change), before
2. Deciding upon, collectively, which (technological) approach or solution for change to focus on in a pilot intervention.

It offers a pragmatic approach for how we propose to further improve the co-creation of knowledge, learning, sharing and translation into practice among practitioners in the energy field. The way the energy system is currently established does not easily permit such a whole-system view which puts human needs, behaviours and (ir)rationalities at the center of interventions geared at system change. Instead, if we look at the energy system through the human lens, we can see that it isn't necessarily a linear relationship starting with supply and ending with the *End User*, but rather a circular relationship which actually starts with the *End User's* need for an energy service. Amongst (rather than sitting above as is usually the way) this view of the system sit the five *Behaviour Changers* (the *Decision-maker, Provider, Expert, Middle Actor* and *the Conscience*, Figure 3).

What is the *Behaviour Changer Framework*?

The *Behaviour Changer Framework* (BCF) is meant to be used as a 'heuristic' to make the mandates and relationships of the *Behaviour Changers* and their interaction with the *End User* clearer. It also enables storytelling for each of the *Behaviour Changers* who are working on a specific behavioural intervention in different domains, contexts and countries.

The "magic carpet", an actual 1.4m^2 piece of cloth, was used in intensive workshops to explore the stories of different *Behaviour Changers* who are working towards a very specific common intervention goal (see combined workshop minutes). The framework was used to explore and visually describe the current situation, different mandates, drivers, barriers, conflicts and intervention tools each *Behaviour Changer* has and their relationships with each other, in addition to their primary stakeholders and the *End User*. It is then used to explore what the system should look like and collectively develop a roadmap

towards a best practice, real-life intervention. Each additional country workshop (up to two workshops per year, per country) explored the changes between BAU and best practice and used the framework to evaluate, reiterate and test completion towards the collectively agreed-upon roadmap.

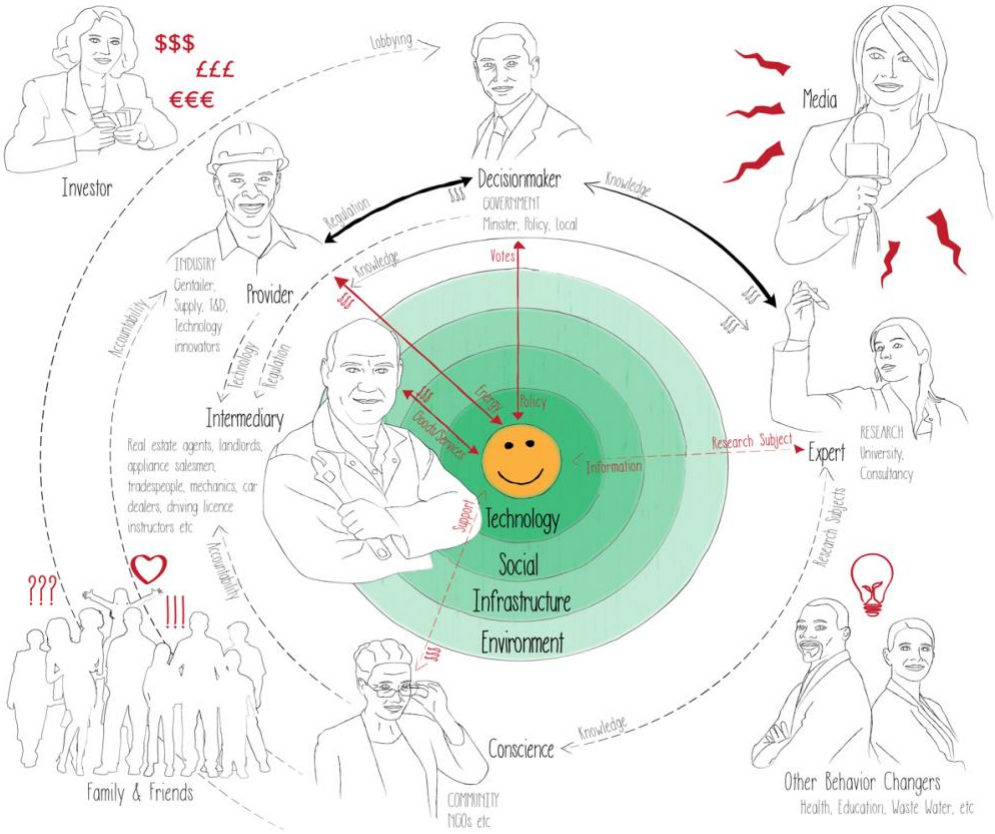

Figure 3. Diagram of the BCF *that works on behavioural interventions on the Energy End User in a generalised Energy System*[3]*.*

The *BCF* thus:

- Acts as a collective impact tool (the process comes before the outcome)
- Helps visualise the energy system through the human lens, showing the current status and barriers, and what is needed in order to achieve a common goal/best practice
- Helps different stakeholders agree on the best possible scenario and then collectively work on solving problems, co-creating the right intervention to change the chosen behaviour(s)
- Helps to evaluate and measure agreed best practice outcomes and how to iterate, if necessary
- Helps identify multiple benefits and how to measure them
- Helps us appreciate each other's world, the lock-ins, restrictions, and relationships, both good and bad which the system throws up

The human actors in the energy system

To be able to change the behaviour of *End Users*, an overview of the social playing field including conflicts and barriers is invaluable knowledge for *Behaviour Changers*. This *BCF* allows an end user perspective with a focus on their behaviour and on the technological and social aspects, infrastructure and wider environment (including political pressures) that need to be changed when solving a complex social problem (Rotmann, 2016a). Next to this end user perspective, a strong focus is given to the *Behaviour Changers* themselves - and their mandates, tools or instruments, restrictions, and stakeholders they need or depend on to perform their role.

The *Behaviour Changers* with often the most 'powerful' impact, the *Decision-makers*, have tools like **policies, taxes and incentives and legislation** to influence behaviour. The second actor-type is the *Provider*, usually focused on providing energy or energy-using technologies. They have different tools,

[3] For a short explanatory video, go here: https://youtu.be/E3A92eFyvNw?list=PLoZ9-YO7tGnoDbnOLmu-cLGC9geztJ0F9

e.g., marketing campaigns, behavioural savings programmes like HERs, or changes to billing systems, with which they can influence *End Users*. The third group, the *Experts*, can develop, validate and criticise technologies and their impact on consumers. Their tools range from scientific papers, (big) data collection and analysis, undertaking interviews, surveys and focus groups in real life or experimental settings. The fourth group is the *Conscience*, usually consisting of non-profit organisations mandated to reduce the social and environmental impacts of the energy system. They use tools like the media, mass marketing and activist campaigns to change behaviour. The last group are the *Middle Actors*, often contractors from a service sector in direct contact with the *End User*. They have behaviour change tools like direct access to consumers, trusted advice, technological information and labels. In addition to various relationships and resource flows (e.g. money for energy or services) between the *End Users* and *Behaviour Changers*, the *Behaviour Changers* also have different relationships of various strengths with one another. Indirect influencers are the *Media, Investors, Family and Friends* and *Other Behaviour Changers*.

Why have two collaboration tools?

The *CIA* is mostly a top-down approach working on the higher levels of social change, whereas the *BCF* can be complementarily used as a way to directly focus on changing the behaviour of *End Users* via a bottom-up approach in collaboration with the relevant *Behaviour Changers,* also enabling a middle-out approach. The *BCF,* thus, offers important additional aspects that should be taken into consideration when creating a collective impact, namely the end user perspective and a clear visualisation of the current energy system, as viewed through the human lens. This includes different conflicts and mandates and different flows of goods and services leading to different strengths in relationships and different tools that each *Behaviour Changer* brings to the table. The *BCF* also includes those who often do not have a direct say in decision-making processes. Incorporating the knowledge about problems that *End Users* experience, the additional bottom-up and middle-out approach and collaboration among *Behaviour Changers*, a "collective" is created which stimulates a feeling of cohesion and empathy. This is a good start for successful communication. Thus, the *BCF* and the *CIA* are able to create a stronger collective impact when combined.

Subtask 9 – Evaluating behaviour change interventions

Beyond kWh, double-loop learning and multiple benefit evaluation tools

When we developed the work plan for Task 24 one of the starting points was the appreciation that DSM projects demonstrate great diversity in goals, scope, participants, resources etc. to match the diversity of *Behaviour Changers'* contexts and needs and their wider environment. As a consequence, developing a generic evaluation and monitoring framework that is widely applicable, yet does justice to this diversity, is very difficult indeed. We realised that finding more appropriate, effective and possibly, validated standardised ways of monitoring, evaluating and learning about successful behavioural DSM implementations was a real and urgent need. Currently, DSM policymakers and other relevant *Behaviour Changers* usually fund and/or support DSM programmes on a rather ad-hoc basis because they lack these means of assessing their impact on contributing towards a more sustainable energy system.

Objectives

- The goal of this research is to develop and validate a set of tools and metrics that can be used consistently for the evaluation of behaviour-based energy programmes including but not limited to eco-feedback, home audits, information and rebate programmes, and social games.
- An in-depth assessment of current (best) practice, cultural and disciplinary idiosyncrasies, country drivers and needs and the best possible international standard (along the lines of psychometric tools like the IQ test - arguably not a perfect indicator of intelligence, but valuable in terms of enabling measurement and comparison).

Deliverables

D 13: An internationally validated set of tools and metrics for evaluating behaviour-based energy programmes 'beyond kWh'.

Beyond kWh evaluation tool

We undertook a review of state-of-the-art research findings, current best practices, and potential standardised ways of how monitoring and evaluating could identify what roles and actions policymakers, investors and other *Behaviour Changers* might take to make behaviour change successful. This review of over 350 residential behaviour change studies published from 2003-2013 was undertaken under the umbrella of the Task by Karlin et al., 2015a ("Methodological Review"). It was found that there is no standardised way of monitoring the impact of behavioural change DSM interventions beyond kWh type indicators (and often even they are not measured in a standardised way): 85 percent of studies did collect some data "beyond kWh," but there was little consistency in the way that these variables were collected or measured. Data on demographics (64 percent), behaviour (62 percent), user experience (58 percent), attitudes (27 percent), and knowledge (21 percent) were collected, but there was significant variation in the questions used within each category.

No standard tool currently exists to conduct such assessment comprehensively and consistently. Such consistency would improve our overall ability to account for variation in treatment effects and verify savings. One of the consequences of not having a bank of standardised and psychometrically-validated survey questions is that research funders lack clear evaluation frameworks to decide on funding practical behaviour change research efforts. Funders therefore continue to rely on the 'easier' technological fixes to our energy problems, and the more common economic or psychological theory-underpinned type of interventions (see also Kallsperger and Rotmann, 2017 for a discussion of the difficulties in measuring and claiming energy savings from behaviour change interventions under the new Austrian Energy Efficiency Law).

The more complex systemic type of interventions that go beyond mere kWh type of outputs thus face severe start-up issues. In order for such a tool to be of maximum usefulness, it will need to be further developed in collaboration across a variety of *Behaviour Changers*, countries/cultures and with input from different research disciplines. This tool was first proposed by Karlin et al. (2015b) and called the "Beyond kWh evaluation tool." The *Beyond kWh* tool was further developed in ST 9 and framed around the New Zealand-led *Energy Cultures*[4] framework. Karlin et al., 2016 state that "*energy behaviour is embedded within the physical and social contexts of daily life; the interplay between behaviour and its contextual influences can be thought of as an 'energy culture.' Behaviour-based energy interventions aim to impact demand through influencing some aspect of energy culture - what people have, think, and/or do. Understanding how a programme does (or doesn't) work requires an understanding of changes in these elements of energy culture.*" The paper presented and tested a set of instruments that evaluate household energy culture before and after an intervention. The tool then underwent further psychometric testing with more than 600 Californian utility consumers (Southern California Edison, 2016).

The tool was then being tested in Ireland for a real-life pilot using public libraries in Dublin as *Middle Actors* to loan out "Energy Saving Kits"[5] (Rotmann and Chapman, 2018a and b). These kits are meant to improve energy literacy and education about people's own household energy consumption and potential infrastructural issues (such as thermal leakage, see also Rotmann, 2018b). We also hope to test this tool on similar pilots in New Zealand and California to show that it is highly adaptable to different cultural contexts, and, thus, universally applicable. It is currently also tested by one CEE sponsors, on a residential behavioural pilot in Vermont. So far, the tool has only been developed for the residential sector. We hope that future iterations will allow us to create modules for e.g. the hospital, commercial office or transport sectors as well.

Double-loop learning

We initiated an expert discussion in 2014 on how a more standardised, practical, robust, generic evaluation and monitoring framework to evaluate both kWh-type of outputs as well as longer-term behavioural outcomes contributing to a more energy-efficient DSM system would look like. We provided a first attempt at initiating and contributing to such a discussion with our second ST 3 deliverable, a "Positioning Paper" (Mourik et al., 2015). In this paper we briefly explain what M&E mean, current M&E practice and how different disciplinary underpinnings of behaviour change interventions influence this. We also discussed the many challenges *Behaviour Changers* currently face when attempting to monitor

[4] http://energycultures.org/

[5] http://www.codema.ie/think-energy-home-hub/what-is-the-home-energy-saving-kit/

and evaluate behavioural change in DSM interventions. These challenges led us to conclude that the traditional quantitative proxies used at present (which are often collected ad hoc and in a non-standard way, see Karlin et al., 2015a) do not correctly reflect if real behavioural changes actually occur. Solely quantitative assessments often miss the details of what exactly is going on, for different people (*End Users* and *Behaviour Changers*) and in different contexts. This is problematic for multiple reasons, and we concluded with proposing an alternative to the current mainstream approach. This alternative includes a focus on *double-loop learning*, allowing for different definitions of success and creating a more participatory approach focused on both process and outcome that makes use of a combination of qualitative and quantitative metrics to evaluate a multitude of parameters for success.

Even though we have not completed a full evaluation 'tool' that can be applied to all possible combinations of interventions in different sectors and domains, we have developed some fact sheets based on the insight that, instead of only undertaking 'single-loop learning', we also need to delve more deeply into the 'double-loop learning' process (see Figure 4 below for explanation). This is especially the case in more systemic, collaborative interventions, as promoted by this Task (after analysis of the case studies in ST 1 & 2 showed how successful such interventions were, compared with siloed, individually-focused, top-down approaches).

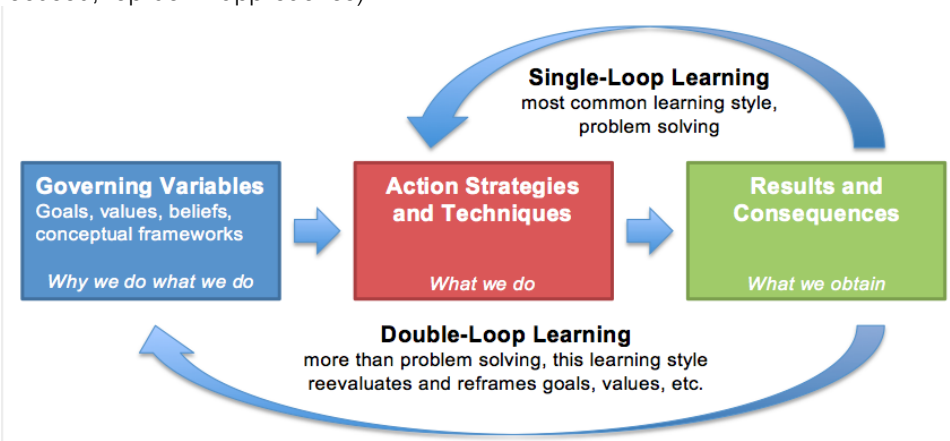

Figure 4. Double- versus single-loop learning. Retrieved from AFS Intercultural Programmes.

In our third ST3 Deliverable (Van Summeren et al., 2015), the factsheet document, we attempted to develop a practical, context-specific monitoring and evaluation template for various DSM tools (which can be used alone or in combination in behavioural interventions), with the specific aim to meet various *Behaviour Changers'* needs for outcome evaluation. This template is developed to match the monitoring and evaluation analysis in ST 1 & 2 of Task 24. The factsheets are a template (completed for three types of intervention tools in the Building Retrofit domain: *Energy Performance Certificates, mass marketing campaigns and subsidy schemes*) which aims at providing indicators, metrics, and ways to monitor and evaluate long-term, identifiable and/or measurable behaviour change outcomes of DSM programmes. These indicators aim to be context-sensitive and contingent on the sector/goals/target groups of behaviour change interventions.

Multiple benefit evaluation

In order to prove ongoing success of behaviour change outcomes leading not only to energy savings, but also health, societal, and environmental benefits such as community engagement or increased species diversity, we also need to look at the additional benefits of behavioural DSM interventions. The multiple benefits of energy efficiency are outlined, with examples, in IEA (2014).

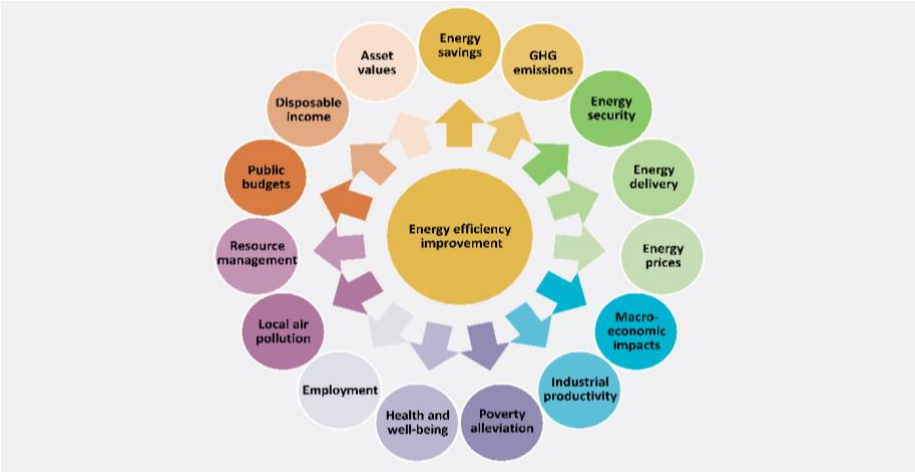

Figure 5. The multiple benefits of energy efficiency improvements. From IEA (2014).

The success of an intervention is usually evaluated on the basis of its cost-effectiveness or its kWh savings (which are often modelled, not measured). However, this does not provide insights about whether or not long-term behavioural change is achieved. Cost-effectiveness and kWh reduction may also fail to capture many of the potential social welfare outcomes and/or impacts such as job creation, positive health effects, reduced environmental externalities etc. Moreover, interventions may have positive spill over effects that not only influence the target *End User* group (e.g. neighbouring effect) but have larger systemic impact, and longer-term effects. Two different types of spill over might be of particular interest, namely spill over to:

- Other people, e.g., peers, neighbours, family and friends; and
- Other types of energy-related behaviour

In addition, energy end users often value other features beside cost reductions which are not included in these cost-benefit calculations (e.g. health or safety improvements). This demonstrates that evaluating success of an intervention should allow the identification of multiple definitions of success – by the *End User* the intervention is targeted at, and the *Behaviour Changers* who helped co-create it. It is, thus, considered valuable in large national programmes such as insulation subsidy schemes, to do some pre-testing of what outcomes would mean a successful programme and to whom (e.g. New Zealand's *Warm Up New Zealand: Heat Smart* programme, see Mourik and Rotmann, 2013; IEA, 2014).

Figure 6. Example of multiple benefits in the transport sector (Austrian report, Kallsperger & Rotmann, 2017).

Of course, a problem with focusing on multiple benefits for different *Behaviour Changers* also leads to the question of weighing up the different (perceived) outcomes. In interventions that take a more comprehensive or systemic approach from the onset, with participation of multiple stakeholders, the

whole process of aligning all these interests and needs becomes a challenge in itself. A solid understanding of where the different *Behaviour Changers* in such a systemic intervention sit in terms of their perceptions of successful outcomes, will help design interventions and their M&E regimes better from the outset. A *CIA*, as used here, can go a long way to aid collecting and analysing these different mandates, drivers, needs and perceptions from the outset. We have thus collected the multiple benefits each *Behaviour Changer* perceived as part of the *BCF* exercise in Task 24 workshops (see e.g. Figure 6 above for multiple benefits from mobility-sharing platforms, Workshop 2 in Graz, September 2017).

Subtask 6 – Understanding the main DSM issues

Background

As part of ST 2 & 4 of Task 24[6], many DSM stories and issues were being identified that lack in-depth understanding and are in need of further research to account for context specificities. Most countries have not clearly identified these top questions with the input from the whole range of *Behaviour Changers*. We acknowledge that the priorities differ between countries, due to different national contexts. We have ascertained and will highlight these country differences (in ST 10). The focus in each country is on three overall priority areas which is then further narrowed down to the top DSM priority that the relevant *Behaviour Changers* (ST 7) will be selected for. This decision-making process of focusing onto top DSM priority areas, collaboratively, is already an important step to foster engagement, empathy with multiple stakeholders and builds on the *CIA* (see above). Collating the relevant group of *Behaviour Changers* from all five sectors for the top priority area in each country enables shared learnings and the co-creation of more focused intervention approaches and real-life case studies according to each of their insights (ST 8 & 11).

Objectives

- Develop lists of top three DSM implementable issues and their potentials in each country.
- Use the *CIA* and the Task 24 Expert Platform to research and review current approaches and practices, nationally and internationally, on these top issues and provide feedback from the different disciplinary perspectives (ST 7).
- Feed these cases, and the ones analysed in ST 1 & 2 into a Toolbox of Interventions (ST 8).

Deliverables

- D 8: List of top three DSM issues for each participating country.
- D 9: Continued collection of case studies and stories (ST 1 & 8).

Subtask 7 – Who are the relevant *Behaviour Changers*?

Background

In addition to the ST 5 expert platform, we have developed more focused networks in the participating countries. The National Experts are coordinating this second layer of country experts. In the US, we have focused on one main DSM topic, namely **evaluating behavioural programmes.**

Objectives

- Identify, with the help of the ExCo and National Experts the most appropriate *Behaviour Changers* focusing on at least one of the top three DSM issues chosen by each participating country.
- Collect detailed information on their specific interests, organisations and roles.
- Use the *CIA* to initiate discussions between different disciplinary perspectives and sectoral contexts. An explicit focus will be on deepening the understanding of the political-institutional context *Behaviour Changers* are operating in and what it means for their capacity to take a more systemic approach to behavioural change.
- Develop national *Behaviour Changer* dialogues in each participating country by holding (bi)annual workshops (ST 6 & 8) to foster mutual engagement, collaboration and shared learning and enable them to build relationships on neutral, trusted ground.

[6] www.ieadsm.org/task/task-24-phase-1/

- Backbone support to set a common agenda, measurement systems, mutually reinforcing activities and ongoing communication between the *Behaviour Changers*.
- Evaluate *Behaviour Changers'* impressions on the effectiveness of the *Collective Impact Approach* and use of narratives as a common language to overcome barriers.
- Collect examples of successful matchmaking stories.

Deliverables

- D 10: National networks of *Behaviour Changers* from all five sectors (government, industry, research, the third and service sectors) in at least one of the top three DSM focus areas (chosen in ST 6); including workshop reports, videos, presentations, stories, blogs, Wiki, etc.
- D 11: Evaluation Report based on stakeholder analyses on the effectiveness of the *CIA* and use of narratives as a common language to overcome barriers.

Appendix 2. Quotes from semi-structured CEE sponsor interviews

What constitutes a behavioural programme?

Sponsor C: "Any definitions are pretty loose with the Commission, so we don't really have any limitations on types of behavioural programmes we can offer beyond them needing to be cost-effective."

Sponsor E: "We actually do programmes that focus on education, purchasing behaviours, etc. in the residential space but we don't regard them as behavioural. The only behavioural savings programmes we claim are SEMs in the industrial and commercial space. These are the only programmes that we think change energy use behaviour versus purchasing behaviour or deeming widgets."

Sponsor F: "Organisationally, we increasingly acknowledge, in terms of programme design and budgeting, that behavioural programmes are broader. Big B behaviour is very much around everything. Marketing programmes are designed with behavioural intentions in the middle. We claim behavioural savings programmes (that are not driven around widgets) on commercial and industrial SEMs. We are looking at how to improve on them with pilots. In the residential sector, we discontinued HERs and are piloting enhanced hardware in the home to drive measurable savings reductions. I actually helped facilitate internal trainings to arm staff on basic understanding what behaviour is and how it works. We are now overhauling our total behavioural project base. We need to do better, and behaviour is the way forward."

Sponsor G: "We need to think carefully about the various behaviours we want to address and design specific approaches for each behaviour. What do we want to accomplish with specific behaviours? Would we like to move towards the desired end state, e.g. in-house temp of 21C? So, what are all the activities that can contribute to this? When looking at all the different thermostats, all with different instructions (manual, programmable, smart) ... the motivations for why people would use them vary wildly. Looking at the barriers to the behaviours, why people do not engage in them, is much more interesting. Each barrier justifies a separate approach. Providing information can be useful to lower the barriers to some behaviours, but definitely not all. Sometimes another solution is required, e.g. a friendly reminder. Utilities sometimes over-tax people with too much information. We need to find proper ways to either automate certain behaviours or remind people in non-intrusive ways. Sometimes it requires reminders, sometimes education, sometimes behavioural economics nudges like framing, social norms or biases."

Sponsor H: "We consider behavioural programmes as those which require no capital investment. I recognize that our definition of "behavioural" may be different to others, e.g. installing more energy efficient equipment isn't regarded as changing behaviour. In this case, the energy reduction comes in the form of more efficient equipment, rather than how that equipment is operated. We mostly focused on creating individual, habitual, behaviour change using behavioural science techniques to change how energy is consumed and equipment used. Examples are in the residential, SMB, and large business customer space. Single family owner-occupied residents - we do a comparison-type programme, neighbour to neighbour HERs. We also piloted a similar commercial programme to HERS, but for 30,000 SMBs. For large business customers, we developed fun campaigns and programmes to motivate both building occupants and operators to change the way they used energy. This included equipment maintenance and operation and even challenging occupants to turn off monitors, lights and to use stairs rather than elevators. We even engaged students in a local University to help businesses identify ways to operate more efficiently. Students benefited from real world engagement and businesses appreciated the additional resources."

Sponsor J: "We do many informational kinds of programmes, but the other more incentive-based ones also have a behavioural component, especially around appliances. Other efforts outside of the traditional list of programmes we report to the Energy Commission, are more around responding to different organisational efforts, e.g. we are in the process of transitioning to time-of-use (day) rates, so different parts of the organisation did customer research on how they were using energy and which customers might be the most critically in need in order to provide them with information to participate in various

incentive programmes. This is one set of efforts that most reflects Task 24 behavioural interventions and case studies."

What constitutes other energy efficiency or demand response programmes that could be called "behavioural" under the Task 24 definition?

Sponsor A: "Typical ways energy efficiency programme savings are claimed are deemed saving values which are based on average estimate savings backed up by engineering estimates and substantiated in written documents called work papers. For example, LED light bulbs: lab studies show how much energy they use in a particular type of fixture and operating environment and compare that with incandescent light bulbs. Then this gets substantiated in a work paper (scientific document) and that is how a deemed estimate gets established. Another way is a custom-calculated approach. There, you have an engineer make calculations of energy savings-based on a model that simulates operating conditions. The third measure is NMEC."

Sponsor B: "We use work papers for other energy efficiency programmes, that's regarded as the "normal" methodology. Another is calculated against a pre-existing base line. The calculated methodology takes many years and is quite expensive. However, the incentives are very high though."

Sponsor C: "The methodology is somewhat similar to our behavioural savings programmes, there is a lot of billing analysis on the residential side, done in-house. The commercial side is usually all third-party evaluation and methods vary by type of measure (e.g. custom versus standard). We also have meter-based savings, get monthly utility data to do the billing analysis. Other pilots we do pre / post metering on site."

Sponsor E: "We use deemed numbers from regional technical fora. In some places, the Power Authority has conservation targets they can claim against. There is a lot of regional oversight. We pay a little extra fee and then it gets returned when they need to claim conservation targets - they usually do more than what is requested. They have their own manual and numbers."

Sponsor F: "Behaviour is around measured changes in energy usage whereas these other programmes are counting widgets. We are looking at hybridising them for heat pumps to support the quantitative data on widget numbers with behavioural insights (i.e. if people use and maintain them correctly)."

Definition by Sponsor F: "Deemed means a document that is agreed upon that deems a specific number for a specific widget; claiming is based on a process that may include more than one project and can include behavioural measures and NMEC."

Sponsor G: "Where does behaviour start and end? Our lighting programme is claiming energy savings up until the install, and after the light bulb is screwed in, it becomes behaviour. This artificial delineation is not always practical."

Sponsor I: "Most other energy efficiency programmes (other than HERs) are part of DSM programmes, those savings are prescriptive. There is research done on different technologies in residential and commercial markets. Each has prescribed savings (with a lifetime savings value)."

Sponsor J: "I feel that there needs to be a marrying of the two kind of analysis techniques - behavioural data with an engineering type of approach. We are talking with companies with exhaustive information on all customers in our service territory (education, salary, home values etc.). They are proposing research to find out who is most likely to adopt PV or EVs, so they can be marketed to - which I find quite creepy, to be honest! Commercial companies who are trying to get the in-depth intel on our buying behaviour - they see us all as consumers. Third party evaluators talk a good game but are not always so good on delivering!"

What is regarded as a credible behavioural programme, in terms of how it is evaluated and claimed?

Sponsor A: "In California, there are two ways: traditionally, only programmes with RCTs and comparative usage programmes, such as HERs could be claimed. There was a lot of effort to change the description.

An AB802 assembly bill has recently passed which allows claiming for savings at the meter, i.e. normalised pre-post energy consumption as observed at the meter. This allows an existing conditions baseline that can show you if an intervention that combines traditional whole house measures (like insulation) along with traditional operational and behavioural measures (like feedback, smart thermostats, timers, etc.) has performed. There is now a combination of methods an implementer may want to add to a package; as long as saving kWh can be detected at the meter, they can claim the savings. This is called NMEC. Implementers can now get paid for energy savings that are seen at the meter, including "behavioural (versus operational) retro-commissioning (BRO). Using savings that can be observed at the meter rather than from rigorous RCTs means a new definition of behaviour may not be needed. The appetite for changing the behavioural definition may have cooled, instead the regulators are simply focused on deeper interventions. BRO measures were not allowed to be deemed or custom calculated as everything was based on widgets and hardware. Widgets are tangible, you can count and verify their installation, which is reassuring to an engineer or inspector - that there actually was a change made. Behaviours are loose and fancy and it is hard to know how long they last. There is definitely a dominance of physical over social scientists in our field."

Sponsor C: "How we evaluate can vary a little, for SEMs there is a model built for each site to model their energy use which includes specificities. They learn from their peers, the first year is engagement. We pay them first year savings, almost like pay-for-performance (or an energy service company (ESCo), although customers are directing what activities are done). We used RCT for HERs, and a randomised encouragement design for Seasonal Savings programmes. We used an RCT initially for smart thermostat programme. Usually, we look at the programme and try and find the best way to identify savings. We use a mix of in-house and third-party evaluators. SEM models are developed by programme staff, then the evaluation is done by a third party. We usually are pretty involved with developing the methods and specifying what we think is best. Our internal policy is to evaluate all pilots, which all behavioural interventions are and then we impact evaluate programmes but only every year to couple of years. We take it for granted but in comparison we do so much more evaluation than most others. Even though we only spend about 2.5 percent of our budget on EM&V - this is similar to other organisations who do less. It certainly helps to have a lot in-house expertise."

Sponsor D: "We use RCTs with our HERs programme, but we use different methodologies to evaluate our App and other flagship programmes - a matched control group (matched on consumption data) for the App and a randomised encouragement design (done by our third-party evaluator). Both showed much greater electricity savings than HERs."

Sponsor E: "As a public utility we have quite different rules, we don't really demonstrate our savings to anyone. There is no regulatory commission, but we have conservation targets we have to reach. We can claim them on the regional numbers and do conservation types assessments. Most of our evaluation is in-house. We certainly check the numbers on any pilot programmes. Behavioural programmes are different to other energy efficiency programmes, there are no deemed measures. HERs would be the closest behavioural savings programme that needs evaluation. SEMs are different, there is not really an evaluation, but more of a verification of sorts. Providing rebates to participants based on how much energy they are saving in the first place, is built into the programme, e.g. we have a 2 percent saving target."

Sponsor F: "We model home by home base historical usage to see if we can measure savings and account for anomalies such as equipment changes. This is not an RCT but modelling and tracking changes to account for double-counting - it includes qualitative analysis and uses an internet-based application and email engagement plan, which includes behavioural questions and has a high return rate.

We also used a randomised encouragement design within buildings with a baseline survey and goal setting. People were told whether they were saving energy or not and had to commit in time how much. It was replicated in a few buildings, they could see their code name on a list in public and if they saved energy in the end they got into a raffle for three months rent (everyone who entered) and another three months rent if they saved energy. It was all based on normalised weather, historical data etc. Those results are informing future programme design. And we got around 10 percent savings! Ultimately, all programmes have to go through savings verification, new programmes have to go through interactions with third party evaluators - these are typically paid by the regulator but they haven't specified hard or

fast rules or guidelines, just good evidentiary baseline. That's why we use both, quantitative and qualitative measures to evaluate our behavioural programmes."

Sponsor G: "Our programmes are designed with measurability and evaluation perspectives in mind. Any programme must be impact-evaluated at least once every six years, whichever methodology is approved and consented by the evaluation committee (including external subject matter experts) who sign off on the final report before it is publicly released. Programme participant times energy savings per unit are the energy savings we report. Impact evaluation can override and adjust any of these findings and these results are often used to update programme assumption values. For upcoming impact evaluation, we will do a comparative study, to see what the most used methodologies are to learn from best practices.

We followed double deflation methodology for our flagship programme where the treatment group is matched to a non-treatment group and measured over time. In pre-treatment, consumption is usually close to equal and then it starts deviating post-treatment. We also use a considerable participant survey which is largely self-reported but tens of thousands of completed responses show a high statistical significance and significant savings."

Sponsor I: "We haven't explored too much what the limitations are yet. We initially filed to do a HERs programme in 2015 but got denied by the energy board as they didn't find it cost-effective for two reasons: 1) Opower is expensive and 2) they couldn't find enough evidence at the time to claim persistence savings. The Board found savings too low for the money they were spending. Another big limitation is that Ministry of Energy funding ends next year, so we are looking to file HERs as a DSM programme (our new government supports DSM but not behaviour). We need to prove programmes are effective seeing that we spend ratepayer dollars."

Sponsor J: "I wonder what's it like elsewhere in the world? I assume it is less hyper-capitalistic than in the US? There can be a lot of distrust in the US towards capitalist claims - and this may lead to more rigorous frameworks and regulation? Bad actors are also definitely an issue."

Are there any regulatory barriers or issues for regulators and / or Implementers?

Sponsor A: "We already have a very complicated regulatory environment. But now the regulator ordered to have all implementation go to third parties. There are new mechanics of approving them and what kinds of programmes they will propose (reasonable, cost-effective, and savings estimates that are achievable) and then do pay-for-performance and find a way to pay them promptly. What we see at the meter will not be too different than what they see much later once their programme gets evaluated and may include exogenous changes. We are hoping that there isn't too much of a difference or that there won't be perverse outcomes where implementers get underpaid and people use more energy. We are claiming against the goals imposed by regulators."

Sponsor B: "What concerns me is if third party implementers believe something is behaviour, but they don't understand how it can be claimed and evaluated for savings. There were a lot of questions from the regulator which needed to be answered first and it took a three-year approval to finally claim the savings from HERs. Having NMEC may open more possibilities but it is also harder as there are more requirements than for RCTs and you have to have customer's usage data - and utilities are really cautious in sharing that data with third parties. What is a new problem for regulators and Implementers is developing rules for how NMEC-based programmes need to be evaluated. RCT rules have been worked out over time, and it took a few years to establish them. We can expect the same for NMEC but they will probably be a bit trickier than RCTs. The model for RCTs is for aggregate programmes, but NMECs aren't run as such rather as quasi-experimental designs. Most NMEC programmes will be opt-in, not opt-out so they also won't have the statistical power and sample sizes of RCTs. However, they will be much stronger interventions though. The effect sizes will be much larger - probably an order of magnitude larger. Guidance to implementers from the regulator is to shoot for minimum of 15 percent savings per household. We are hoping for at least 10 percent. We first need to get meter signal above the noise to prove that it works. Once regulators are not familiar with concepts anymore you need to show how to validate savings, how to claim them, where they originate from, etc. These are the biggest challenges when doing any behaviour intervention. Couldn't we just use the AB802 guidelines? Everything we usually did was based on the widgets we sold or shipped to customers, which is so much more straightforward. HERs have a lot more probability and bias potentially which needs to be sold (i.e.

every new intervention has to go through the same process as everyone wants to avoid taking on what is perceived as high risk). Even though they spent the money and saw the results, they still felt unsure and uncomfortable in changing tack. We are trying to develop a new way of testing HERs that can reduce the bias and want to do a White Paper for the regulator. Now, we are only allowed one year of savings claimed for behaviour, but we want to get two to three years of savings. One of the variables is useful life, which right now is only one year (i.e. the measure only effective for one year). We want to change that."

Sponsor C: "We do not have the same regulatory restrictions as California, we have a pretty collegial relationship with our Commission. In California, they seem to be very much in the detail as to the measures that get developed and how they get measured. Here, the Commission gets regular reports, but they don't dig in and question us. It seems enough that we conduct the interventions. It also helps that we are not a utility, but a third-party non-profit that administers programmes which aren't reimbursed the same way a utility would be. Everything we do needs to be cost effective and measurable on a total resource cost - so anything that comes out to be cost-effective, we can pursue. The flip-side is that we can't do certain things that a utility can do like bill alerts, switching people between tariffs etc. We can't do things just for customer goodwill. However, we'd like to include NEBs in cost-effectiveness, and this is hard unless they are significant and quantifiable. Our Commission will give exceptions for some NEBs, but it requires a process which is much more complicated than for kWh. We would like to pursue things with more health and water benefits, but these may take more money to quantify them. On intra-organisational barriers, we have many creative people in the organisation but are too restricted to short-term views on savings instead of being able to develop longer-term behavioural programmes. We would need the establishment of different goals beyond single year savings. It's certainly possible to establish longer term goals, especially in terms of climate change, health and social benefits, and also a more holistic view towards energy efficiency and RE. But our leadership is reluctant to do anything that makes them a potential political target."

Sponsor D: "We have a cap on what we can claim for behaviour savings, ramping down to 15 percent by 2022 (from originally 24 percent). They range from 15-20 percent on the gas side. This is due to regulatory barriers - they were agreed on the last plan filing. There is an issue around stakeholder perception of behaviour savings, it has low measure life in our State (only one year), which results in low measurement counts and it impacts on the cost-effectiveness of the programmes. It is also hard to quantify what people are doing. This could be due to a variety of things e.g. buying more energy efficient appliances (they account for double counting but if people don't claim the rebate you simply can't know!"

Sponsor E: "We don't really have any regulatory barriers, programmes are either verified by us or a deemed measure. We don't really have restrictions. If we develop a novel programme we just have to evaluate and document any savings. We have so much more flexibility than IOUs."

Sponsor F: "We are feeling the pain of budget cuts on exact areas like the hard-to-reach pilot - we had to put out earlier fires so the evaluation budget was cut, even though it was initially included. We now have evaluation budget only for things that are critically important, like really big bids. Smaller side pilots might just test one little piece, and we don't take it to the regulatory bank but learn from insights for future programme design. Even though 50 percent of the time we do EM&V in-house, we do have a third-party system and a regulatory body. In that environment where everyone (including third party vendors) suffer budget cuts, how do you deal with questions around oversight? Will the regulators simply conclude it's too expensive to measure and thus not include more novel programmes? There's definitely a challenge of measurement and budget - we want to avoid what happened in California with too narrowly focused claims on RCTs. Mass-customised NMEC paired with good quality surveys would be a good change. Budgetary constraints and engagement with regulators are issues. We would love to codesign evaluation methodologies with them to capture their risks and concerns. Right now, they are over-collecting data to try and ensure it meets their needs. As long as conversations between utilities and regulators are trusted it's all good but at times they are stretched so thin and lack expertise sometimes, then it causes a big log jam. Behaviour is showing up on a demand resource plan here (which informs savings targets and goals), whereas often it doesn't show up as it's not a widget programme. That is a problem in terms of behavioural savings potential, which they think is only 1-2 percent, when we know it's more like 30 percent."

Sponsor G: "The DSM portfolio aims to meet the DSM plan targets. We also have challenges with strict

privacy rules and regulations, it is extremely difficult to upsell or cross-sell other programmes to customers and as a result, most programmes recruit participants separately. It is a programme focused approach, rather than a customer focused approach. As a utility we should be looking at working towards a model where we look at it as an end perspective, with a more customer and end use focus than being so product focused. What is the regulators' hesitancy of not embracing behavioural programmes? The solution is not looking at cost or potential but validity by looking at impact evaluation."

Sponsor H: "There is a preference to maintain status quo. Changing from existing methodologies is challenging. Generally, new methodologies may require a lot a convincing. Internally, we need to clearly demonstrate that the savings really exist (and make sure to avoid double counting) There is definitely a big difference between public and private utilities. What if your mandate is not the overall societal good if you are an IOU? Public policy in the US is set at the federal level and very politicised. We need to develop business models to solve public policies that are mandated. The Energy Policy Acts put forward in 2007 were billed to meet certain standards, e.g. climate change standards. Public policy is set and business models will change accordingly, but if you are heavily-regulated you can only do so much. We need regulatory approval, we tried to claim savings from our behavioural programmes and could do from pilots but the second round of evaluations showed the savings were less. So we had to cancel these programmes. When you start targeting smaller businesses, it can erode your performance. Smaller buildings don't have the same opportunity for savings as they just don't have the same consumption."

Sponsor I: "Proving not to waste ratepayer money is hard. Political change causes issues too - DSM is now the only energy efficiency model this government accepts. It is hard to file anything for approval that doesn't show high enough savings or is experimental. It needs to be proven in the market and has evaluated results and we then need to show we can do something similar. But there is an opportunity to test out pilots."

How can some of these barriers be overcome?

Sponsor A: "I think a movement, first toward adoption of national evaluation standards within the US, such as those championed by the Uniform Methods Project, and then by adoption of such standards internationally, will go a long way in overcoming these challenges."

Sponsor B: "I think the 2017 Draft White Paper would have been very useful. How do you calculate something new or different and ensure people adopt it and accept it so you can claim savings as actuals? How do other countries prove that their goals are realistic and actual? We also tried to push for NEBs, but the regulators have excluded it. Ultimately, we may need modelled estimates, as claimed or calculated savings are too hard to prove for new interventions."

Sponsor C: "We really need programmes that have already been tested elsewhere and shown to be promising. We can't take the risk as first movers."

Sponsor D: "We have meetings with staff to address their concerns on an ongoing basis. We also have stakeholder meetings and are trying to get qualified consumer data to backup any behavioural changes we claim."

Sponsor E: "Start small and pilot it, show proof of concept."

Sponsor F: "Have a good, trusted relationship with the regulator. Also, fluency with data science as an organisation is required, and trust in data and statistics. It is important to have the people in-house who can crunch the data so we can trust the data. Broadly, we need to measure everything, and expect that there are a lot of behavioural savings out there. We need to grow the behavioural potential from 3 percent to 30 percent and put it on the demand resource plan. We need to overcome perception issues on the smallness of savings by showing with better measurements and actual usage changes that it is much more. That includes triangulation of big data with "little data" from qualitative surveys. I would like to have someone put a big number on the price of waste or cost that arises from US utilities underperforming in the behavioural space. There is a cost of behaviour not being reported to the level it should be. When you then layer NEBs on top of it, you can show that it is a huge cost to ignore."

Sponsor G: "We were one of first to claim behaviour savings. Solid impact evaluations have proven that

the behavioural energy savings are real. For new behavioural concepts small scale pilots are conducted first, show proof of concept, show people that it works with just a few hundred customers, get very tangible results, and then upscale. Many people find it difficult to imagine what to picture and the small-scale trials allow us to tell the story. You also need to pick your battles - it depends on the project and situation and there is always trading and negotiating. It is very difficult to start multi-million-dollar projects because the climate just isn't right."

Sponsor H: "It is interesting what other countries are doing. Providing different insights is about legacy and ongoing, long-term changes as well. So it is worthwhile to make this comparison even if we are comparing apples and oranges because we are on a path of some convergence. We work closely with many different organisations, there are continued conversations and demonstrations of benefits. What we are doing here with Task 24 is really worthwhile. It would be very hard to accomplish individually but with the Task 24 leadership and CEE. It will be really interesting to see everything from a more global perspective."

Sponsor I: "If we were to file for HERs, it's down to showing that there are persistence savings that are independently verified to the regulatory body. We have to show that they save more than we initially filed in 2015 to get them to approve it. But I think it's better than we initially thought as annual savings are quite a bit higher than we forecast (we had a colder winter than usual). Right now, they are three times greater than we forecast, so it will be an easier sell next time."

What constitutes a persistent behavioural programme and how is persistence usually measured?

Sponsor A: "Yes, we conducted persistence studies with HERs reports. We sent out the reports for a couple of years, then stopped sending them and observed how long savings continued to persist in the absence of the reports. We looked at average endurance of savings and found that they do persist, but it is complicated as you need to choose a term of how long they last. It's easier when you look at half-life with energy efficiency programmes - i.e. how long does it take for e.g. half of the LEDs burn out - that is the effective use-for-life (for widgets). With HERs - how long does it take for half of the savings realised from the HERs cohort to diminish? We found around three and a half years. But it is very hard to distinguish it from, for example, the purchasing more energy-using equipment or making other changes in their household. We do compare them to control groups. It all shows very low ongoing savings, within the standard error (1-3 percent), but there are still large savings in aggregate. You need a very large cohort to see any savings - and ultimately, the customer pays Oracle for these home energy reports."

Sponsor E: "We don't really claim or study persistence. We don't look at how many homes had equipment replaced that is still in use, for example. Our teams measure estimates of, for example, shower heads where they rely on available data on persistence and do take it into account in savings estimates, but that's about it."

Sponsor F: "We looked at it with our multi-family's pilot but it was inconclusive and probably not long enough. Persistence is a misnomer, it basically is an artifact of HERs but broadly, if we are shifting more towards NMEC we can always see if there is lasting change. Most people want a deemed number for persistence whereas we should look at how big a change we can get and how we can make it bigger over time."

Sponsor G: "We are heavily interested in the topic, we haven't done such a study yet but now that we have nine to ten years of participant data available. Our upcoming impact evaluation will have a persistence study attached to it."

Sponsor I: "Our one behavioural savings programme is still new, we will run HERs for another heating season (to conclude March 2019). Afterwards, we will look into persistence savings after three years. Other DSM programmes look at lifetime savings, but this is just estimated from technology."

What is regarded as a "hard-to-reach" customer and how are they engaged?

Sponsor A: "We have quotas for reaching hard-to-reach communities. No programmes are specifically targeting hard-to-reach customers but a focus on having a certain proportion of the delivery of our

programmes to reach hard-to-reach areas within the service territory. This is separate from other programmes outside the energy efficiency space that are dedicated to qualifying customers that are lower income. It's a different funding silo. 40 percent of customers qualify for reduced cost service and retrofits of their homes (CARE - Customer Access to Reduced Energy). It is subsidised by customers that pay higher rates. But it is all siloed - there's no unified way of deciding who is eligible, some are done by the federal government, some by the State, all different government entities. Customers are aware in California that in terms of per kWh charges it is amongst the highest in the country. More affluent people live on the coast, thus have lower overall usage. Higher proportion of subsidised people are inland, and they have higher rates of consumption.

We look at multi-family apartment buildings and there are split incentives for multi-families as some complexes are master metered. In some, tenants own the meter and pay the bill, in others it is the landlord. Especially the latter are regarded as hard-to-reach customers. Also, regarded disadvantaged communities as such and use a list of zip codes (Goldbergs hard-to-reach). There are multiple definitions depending on who you talk to, what industry they are in and which programme they are managing. They never mean the same thing."

Sponsor C: "We consider hard-to-reach as rural, non-white and low-to-moderate income customers. We have a diversity/equity/inclusion operations plan, which outlines how we can better address and engage underserved communities."

Sponsor D: "Each group has a different definition of what they regard as hard-to-reach."

Sponsor E: "We have a different definition to others, such as Californian utilities. We think of it as anyone who is hard to get to participate (e.g. we don't know how to contact them). This overlaps with other definitions but there are some distinctions around the difficulty of contacting them. Rental units are hard-to-reach, so we have pilot programmes for "free everything" to get them to upgrade their home (for landlords only). Tenants are even more hard-to-reach. There is also a difference between rental units that are hard-to-reach rather than renters who don't have decision-making power (e.g. retrofitting). Low income is not regarded as hard-to-reach, though traditionally it was. Our low-income residential customers are actually well-served by targeted programmes. It is where we don't have targeted programmes, we find underserved (e.g. rental units) or under-represented (e.g. small retailers and offices) customers."

Sponsor F: "We have a discontinued pilot, we had a programme designed specifically to serve commercial customers like SMBs that are really hard-to-reach and they didn't get served so well. But there were gaps between what the vendor thought they could do and what they could actually do. But we still got interesting qualitative insights - we want to do a variation of that programme next. In the residential sector, is it hard-to-reach or hard-to-serve (really low income or renters or multi-families)? There are several definitions. As result of the difficulty in finding funding for such pilots, in terms of their novelty and hard-to-measure-ness, they often hit a snag and you run low on budget/time and report writing suffers. We often chose not to do the report and just take away insights internally."

Sponsor G: "Hard-to-reach is an inaccurate term. If you can only reach out to some customers by means of snail mail it doesn't mean they are hard-to-reach, it is just through a different means of communication. What others have mentioned as hard-to-reach are target segments that are difficult to recruit or design for. The difficulty is not in reaching them, it just may need more effort or different communications vehicles, but if they get their utility bills they can be reached. I think a better definition may be segments that utilities traditionally struggle to get a decent market share of or uptake."

Sponsor H: "When we think about hard-to-reach, we think about renters or SMBs - i.e. how to get in front of the decision-makers. Not hard-to-reach in terms of a specific aspect of a programme, but how to get the right info to the right person. Account managers know most of the large business decision-makers, but it is harder to understand and reach decision-makers in smaller businesses. Renters can alter how they consume energy, but often property owners need to be involved if capital expenditures are required to reduce energy consumption. There is also a difference between hard-to-reach and hard-to-engage. We send a utility bill to everyone, so presumably there is at least one communication channel in place, but creating engagement and action is more difficult."

Sponsor I: “We don’t have a hard-to-reach programme, the reason is that our Energy Ministry was very vague on their goals with the HERs programme. They just wanted to achieve greatest savings with the allotted funds via Oracle/Opower then, cherry pick the highest users to get the best savings results. We have 140K programme participants and all are the highest users. We were trying to figure out how to define hard-to-reach, but there were different opinions, and no agreement in the end. One definition was whether customers have technology, e.g. rural areas with limited internet (one definition was around delivery channels). We are mostly in rural areas, so these are the most common types of hard-to-reach customers. Other utilities don’t have this problem.”

Sponsor J: “One of our staff has been working with researchers to disaggregate meter data and identify customers who have really high air-conditioning loads, for example. What are characteristics of heating and cooling load, etc. and then specifically look at low income customers and how they could be targeted with energy efficiency and weatherisation, etc. efforts. But this is not a specific programme, and we don’t claim for it.”

Appendix 3. Task 24 Expert Interviews

Short, written interviews were also conducted with three of the national experts of Task 24 (Austria, Sweden and New Zealand). They all operate under quite different regulatory regimes and their answers reflect their views of their country's contexts. These interviews can be used to further contrast how behavioural interventions and energy-savings programmes are progressed in different countries.

From Teresa Kallsperger, Austrian Task 24 Expert

"Energy utilities, who sell more than 25 GWh of energy to energy users in Austria, have a legislated mandate under the Energy Efficiency Act (EEffG): they have to save 0.6 percent of their previous year's energy sales via energy efficiency interventions. This contributes the utility sector's part of Austria's legislated mandate to the EU's EED. The goal of the Austrian EEffG is to reduce the energy end use through the energy efficiency interventions in the utility sector by 159 Petajoules (PJs) by 2020. In addition, Austrian public agencies need to contribute 151 PJs. Of the 0.6 percent of the utilities' annual energy savings, 40 percent needs to be contributed by Austrian households."

"We have a document from 2016 which outlines the legislated interventions in households. From this report you can see which interventions were claimed by the utilities (34 percent were in heating and hot water interventions; 25 percent in retrofitting; 7 percent in lighting; 6 percent in mobility interventions). However, it is unclear which of them were actually accepted by the Monitoring Agency - especially for example, the 30 percent of individual interventions that were not clearly defined in the monitoring document but which claimed to have led to 44 percent of the energy savings (Adensam et al., 2013; see also Kallsperger and Rotmann, 2017; and Task 24 Endbericht, 2017 [in German, summary in the database]). One feedback from utilities however was that it was cheaper for them to pay the penalties for not achieving their targets than it was to implement energy efficiency interventions."

Mehmet Bulut, Swedish Task 24 Expert

- How important, if at all, is the role of behaviour in your current energy portfolio programme? Is it specifically mentioned or embedded in any policy, such as energy efficiency, energy transition, climate change etc.?

Behaviour change is regarded as very important to achieve the Swedish climate goals and the government has therefore introduced several incentives in recent years to influence behaviour change. Hourly metering of electricity for example is offered for free and there are obligations imposed on energy suppliers to provide data on electricity use. Another example is the financial support for electric bikes to reduce the use of personal vehicles. There is also growing interest in nudging and inter-disciplinary research that incorporates social sciences, such as behavioural science, design, and anthropology.

- What do you understand constitutes a behavioural savings programme in the European context?

Any programme that encourages energy efficient behaviour, either through uptake of energy efficient products or influencing usage patterns.

- How does this differ from other energy efficiency or DSM programmes that do have a behavioural aspect (e.g. technology purchasing or uptake, technology maintenance, decision to retrofit, ask for an energy audit, etc.)?

It does not differ. Technology procurement, for example, is important for developers and housing companies to install energy efficient appliances or construct buildings that are more energy efficient. This could then be complemented by influencing energy efficient behaviour through energy policy tools that address end users (could be education or carrot or stick approaches).

- What do you regard as credible evaluation methodology of a behavioural savings programme? How does this differ from other energy efficiency or DSM programmes?

I would assume a combination of data (quantitative) measurements and qualitative surveys would constitute a credible evaluation methodology. Similar methods are often used to evaluate energy efficiency or DSM programmes.

- How do you perceive the impact of regulation on behavioural savings programmes?

Regulation plays a role in influencing energy behaviour, but there is a need for services and other measures to keep users interested and engaged. It is therefore important to combine energy with different aspects of living, for example, an internet-based application that shows energy use of different

appliances can provide information on how much the juicer was used (healthy living) or if the kids played too much Playstation that week etc.

- How closely do you work with regulators? Do you perceive any regulatory barriers of behavioural or energy efficiency programmes? Does your organisation have a position on how to overcome such perceived regulatory barriers?
We have ongoing cooperation with the Energy Market Inspectorate ("Inspectorate"). The perceived barriers are mostly related to the policies that are developed by policymaking institutions, such as the Ministry, and not necessarily the regulator. The Swedish Energy Agency reflects its position on the proposals of the Inspectorate or the government through its official responses to reports and proposals.

- Does your organisation have a position or interest in the persistence of behavioural savings programmes?
The Swedish Energy Agency has strong interest in using behavioural aspects to encourage energy efficiency both through everyday energy behaviour but also by influencing purchasers for the dissemination of energy efficient products on the market. This is also why we have supported Task 24 from the beginning and will be part of the next Task on hard-to-reach energy users.

- Does your organisation have a position or interest in how to engage hard-to-reach energy users?
Sweden has the lowest energy poverty among European countries, so energy poverty is not a prominent issue. There are, however, issues surrounding the security of electricity supply, which can be considered a type of energy poverty. Users experiencing electricity cuts, who mostly live in rural areas and are very much dependent on electricity due to the use of heat pumps, could be considered hard-to-reach energy users. The Inspectorate and the Swedish Grid are the main governmental actors that are responsible for the security of electricity supply. However, the Swedish Energy Agency also works to promote the vision of a sustainable energy system which is fully renewable and has high security of supply. We are also particularly interested in creating a more sustainable transportation future, as the Swedish government has a target to be free from fossil fuel imports by 2030.

New Zealand's Utility Obligations

The main government organisation in charge of energy efficiency and conservation efforts in New Zealand is the Energy Efficiency and Conservation Authority (EECA). It releases the *New Zealand Energy Efficiency and Conservation Strategy* (NZEECS) every five to six years, in addition to the Ministry of Business, Innovation and Employment's *New Zealand Energy Strategy* (NZES). The targets under NZEECS are:

1. Decrease in industrial emissions intensity (kg CO_2-e/$ Real GDP) of at least 1 percent per annum on average between 2017 and 2022.
2. Electric vehicles make up 2 percent of the vehicle fleet by the end of 2021.
3. 90 percent of electricity will be generated from renewable sources by 2025 (in an average hydrological year), providing security of supply is maintained.

None of the targets are binding or monitored using third party evaluators.

The electricity regulator is the Electricity Authority (EA). New Zealand is regarded as having one of the most deregulated electricity markets (see Ministry of Business, Innovation, and Employment (MBIE), 2015). Utilities are part of the New Zealand Emissions Trading Scheme but have no direct energy efficiency obligations. Only 0.5 percent of a customer's utility bill is spent on "Market Governance" which includes costs for energy efficiency programmes and to run the EA (see EA, 2018). The EA runs a scheme called "What's my Number" to educate consumers about the benefit of comparing and switching retailers. EECA runs programmes and information campaigns to reduce energy use in the residential, industrial, commercial and transport sector. Main campaigns are the Warm Up New Zealand insulation programme (Mourik and Rotmann, 2013) and various Minimum Energy Performance Standards and Labelling schemes. EECA also has a "Rightware Tool" programme to help consumers compare energy efficiency of different appliances. New retail pricing plans include time-of-use tariffs that let consumers and tools monitor their own consumption online. They can then make savings by shifting their electricity use to off-peak times or take advantage of low night rates to charge their electric vehicles.

Dr. Daniel Gnoth, PowerCo, New Zealand National Expert

- How important, if at all, is the role of behaviour in New Zealand's current energy portfolio programme? Is it specifically mentioned or embedded in any policy, such as energy efficiency, energy transition, climate change etc.?
Doesn't seem to take much of a priority in New Zealand, in the EECA strategy focuses more on information and purchase behaviour but the National Energy Research Institute one seems to shy even further from such initiatives. Behaviour focused programmes also haven't been awarded to any of the large MBIE research funding bids in recent years.

- What do you understand constitutes a behavioural savings programme in the New Zealand utility sector context?
You mean energy savings programmes which leverage or target behaviour change specifically? Well, there are a broad range of options, but ultimately the utility sector by definition needs to leverage all initiatives across the value chain as utilities are 'Middle Actors' which rarely intervene/engage with end consumers directly. So any initiatives at all levels of government, non-profit organisations as well as the private/commercial sector could constitute some form of utility involvement. Many of the behaviour change products and services that are being bundled into retailer offerings and web-based applications will leverage pricing signals and information provided by the utility sector.

- How does this differ from other energy efficiency or DSM programmes that do have a behavioural aspect (e.g. technology purchasing or uptake, technology maintenance, decision to retrofit, or ask for an energy audit, etc.)?
Generally, the initiatives would be pricing- or policy-based on which provide signals and incentives to the market. Sometimes they are information-based (providing research papers etc) that are again usually targeted at market participants to inform and educate.

- What do you regard as credible evaluation methodology of a behavioural savings programme? How does this differ from other energy efficiency or DSM programmes?
Energy savings/reduction or load shifting are generally measured by utilities in demand - kW (not volume - kWh terms). Utilities still require the same evidence as other initiatives, and more granular information about customers and their experiences is just as important to utilities as it is to other participants in the energy system. Ultimately, customer satisfaction is a focus for all businesses and is at the heart of any initiative that is being implemented and measured, in other words, does energy saving actually help the customer, is it what the customer wants, and will the customer be able to keep realising the benefits?

- How do you perceive the impact of regulation on utility-led behavioural savings programmes?
Especially in regulated industries such as network utilities in New Zealand, regulation creates the framework by which industry can engage with their customers. In other words, for businesses that are regulated, regulation has everything to do with how energy savings programmes can be implemented.

- How closely do you work with regulators? Do you perceive any regulatory barriers of behavioural or energy efficiency programmes? Does your organisation have a position on how to overcome such perceived regulatory barriers?
Utilities work closely with regulators to ensure rules are understood and adhered to but also to help provide context and insight around the changing nature of their business. Typically, there aren't any barriers to energy efficiency programmes for regulators apart from those which are inherent in policy design. There aren't any specific barriers to utilities being involved in energy efficiency programmes, it's more about what the role of the utility should be in this space in the eyes of the consumer/market and regulator.

- Does your organisation have a position or interest into researching the persistence of behavioural savings programmes?
Evidence is crucial for any business when making decisions. The perceived certainty that comes through interventions which do not rely on engaging with consumers directly continues to be a popular approach as the persistence of any behaviour change intervention is rarely monitored to much satisfaction.

- Does your organisation have a position or interest in how to engage hard-to-reach energy users?
Depends on the definition, but ultimately all energy providers would want to ensure that their customers feel connected and supported, and that providers are always able to find better ways to better engage their customers.

Appendix 4. International Input from BEHAVE and BECC Conferences

BEHAVE conference, Zürich Switzerland, September 5-7, 2018

In Switzerland, we used audience prompts and prepared handouts, of which 30 were returned by the almost 60 attendees, who hailed from 13 different countries. Summary findings of these responses show that the majority of respondents (23/30) came from the research sector.

To the question **why behaviour change was of interest to them and why they were at the BEHAVE conference**, six expressed that behaviour has an overarching impact on everything and because of that we, as a society, need to change our lifestyles.

Some choice quotes below:
"*Solutions are implemented by and for humans, thus behaviour has an intrinsic but often neglected role in reaching a goal. Research has a way to uncover the potential of behaviour and determine ways to use if for good/effectiveness."* Swiss researcher

"*I work as a communicator and in my organisation, I have the opportunity to communicate with people on energy efficiency and the protection of environment - what they can do and why they should do it. In order to communicate better and understand them better it is crucial to understand behaviour."* Croatian government official

"*My altruistic motive: to contribute to a better world; my hedonic motive: to have an interesting job; gain motives: to be knowledgeable enough to get more research funding."* Swedish researcher

Most attendees came to BEHAVE to learn the latest research (19), network (9), and/or present (6).

"*Because I think technology alone cannot save the planet. We need to change our lifestyles completely to reach our goals. It is THE conference if you are doing research in this area. I'm hoping to learn from other researchers' experience with things like demand response."* Swedish researcher

"*This is exactly my current work! I'm working on energy efficiency (for the military) through a multi-disciplinary approach including behaviour. This helps me to evaluate my project by seeing what peers are doing and giving new ideas*." NATO policymaker

"*I'm here to get an overview of the IEA's energy behaviour project. I will use it to put my own research project in a context. This workshop: to learn how to evaluate the improvement of behaviour change."* Swedish researcher

On the question of **how do you define behaviour / change**, a range of definitions were provided. The most prevalent were:

- social science techniques (practice theory was mentioned five times!) or insights to change behaviour and save energy
- a set of factors that drive energy decisions
- the result of a behavioural economics-informed intervention
- what can be observed from a difference in behaviour but there are many externalities
- lifestyle and cultured change
- behaviour change has evolved into sustainability, which also means an acceptance of it
- energy behaviour improvement instead of behaviour change

"*Well, sometimes the focus is on "purchase decision," "action" or "practice," but "behaviour change" as an aggregate term is fine for me."* Swiss researcher

"*I use your Task 24 definition on energy behaviour change as reference in my work! I came to your Task 24 workshop at BEHAVE 2016 in Portugal."* NATO policymaker

"*Individual behaviour is essential as we all work as individuals no matter what we do. Behaviour change is evolved for sustainability, be true to your values (and improve). For my organisation it is also improved acceptance of sustainability, understanding that the business will not suffer from sustainability. In contrary, it will prosper.*" Finnish researcher who was part of organising all BEHAVE conferences

"*It's the change of habits, much more than just an aspect. It should be lifestyle and practice and culture change, together with the change in general - one society faces.*" Croatian policymaker

"Behaviour is what we can observe. A behaviour change is when we observe a difference in behaviour. There are many other things that we cannot observe. Such as an individual's values, etc. Behaviour is the outcome of an array of internal and external factors." Swedish researcher

"Behavioural change refers to being flexible in response to price signals when using electricity. As a private citizen it entails changing your lifestyle." Norwegian utility

"*I'm here to find out more about behaviour change so I cannot provide a definition, for sure it should be a lifestyle and a cultured change. Our projects use school buildings but we also want to do educational projects. Thus, we don't want to concentrate on the education to an individual's behaviour but also to the behaviour of a group.*" Hawaiian Energy Institute

"*Behaviour change has to do with the energy culture framework definition* (see Stephenson et al., 2010). *Therefore, to produce a change in energy culture you have to work on not only energy practices but also material culture or cognitive culture around energy.*" Spanish researcher

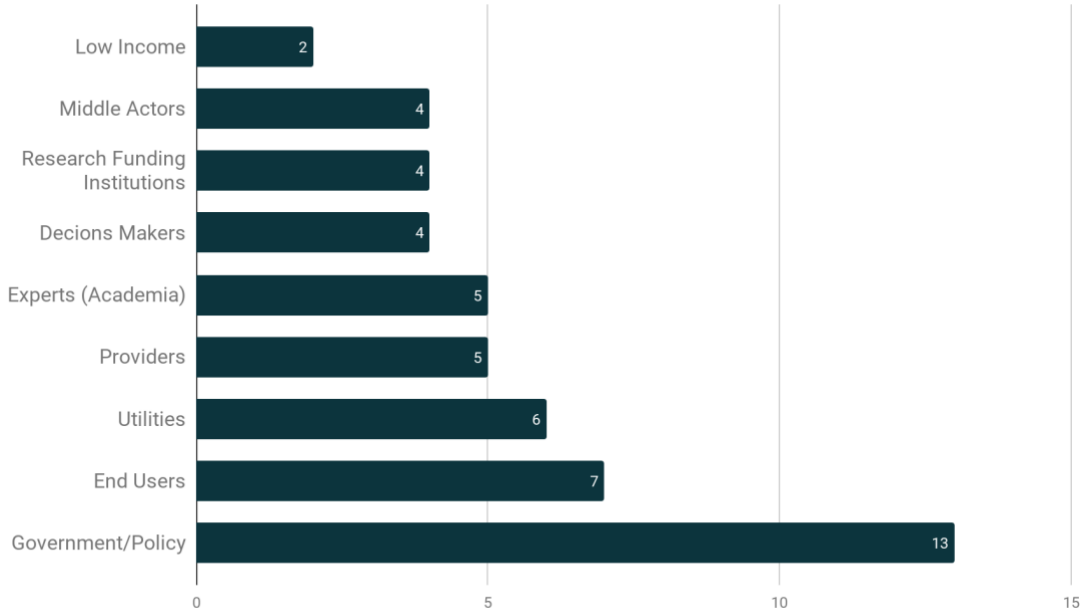

Figure 7. Participants' main stakeholders / audience.

The participants' **main focus or work objective** was as follows:

- applying behavioural economic insights to energy transition and sustainable policy
- better understand the interaction between people and technology
- decarbonise the European commercial and residential sectors
- design information and communication technology (ICT) interventions
- end user education
- identify household mechanisms that can address their energy needs effectively, efficiently, sustainably and in a culturally-appropriate manner
- identify why there is a gap between what people do and what they say they will do
- implement and evaluate interventions targeting decreasing energy use
- user satisfaction

- research how people interact with technology
- how to engage consumers
- to support the use of energy efficient technology that provides environmental benefits without interrupting the everyday life of the user

"*I have not used any particular approach, except for conducting surveys and running programmes to measure the effect of different variable (factors) on energy use (e.g. income, demographics, etc. effect the amount of biomass use).*" Central Asian researcher

"*I use an experimental approach to establish a causal link between a treatment (behaviourally-informed intervention) and a desired outcome (change in behaviour); and behavioural economics research to inform interventions in sustainability policy area.*" EURAC renewable energy institute researcher

We asked **what is your preferred approach to behaviour change?**

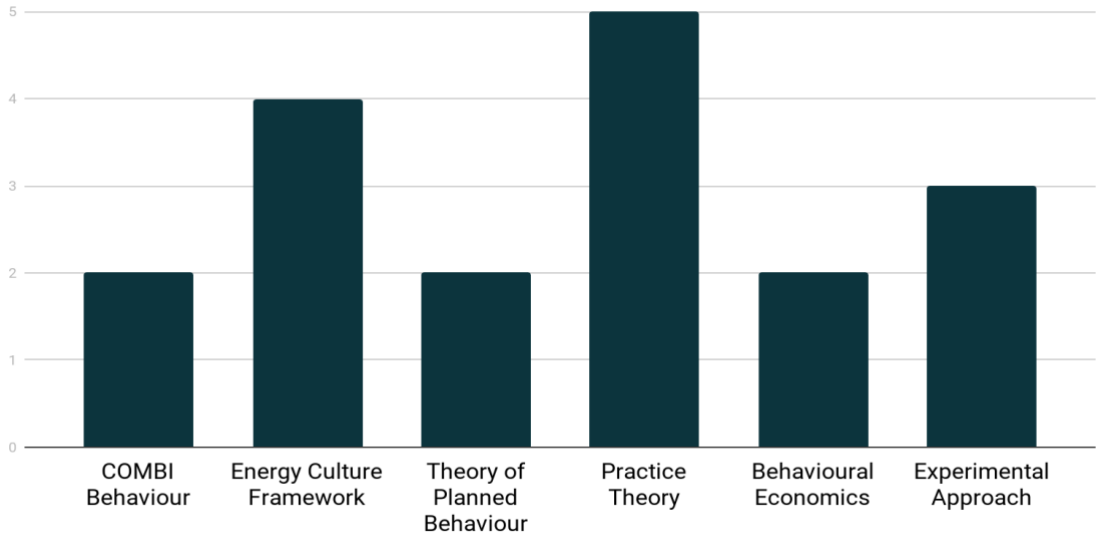

Figure 8. The main models and disciplines used by the participants.

A majority of participants (17) agreed that **behaviour change suffered from a credibility problem**.
Many said the credibility issue is related to funding (x4) and that the social sciences disciplines are underrepresented in funding (x3). Other issues included:

- can't be sure the intervention was the cause of the behavioural change
- it's difficult to prove energy savings (x4)
- lack of consistency and replicability of the projects (x2)
- lack of trust from engineering colleagues (x3)

"*I haven't faced directly credibility/ double-counting problems. However, it's more than evident that in the number of calls or volume of research funding, that the social sciences are hugely underrepresented compared to technology and business models.*" Swiss researcher

"*The credibility problem comes from the lack of trust from engineering colleagues. Doing user interface and behaviour change work, I have been told 'we don't care about how people feel, what about the numbers?' Too often, qualitative work just isn't taken seriously.*" Polish researcher

"*Yes! I work predominantly with male military engineers who I often feel look down on me and my social science colleagues/peers. The only time I feel they give women credibility is when their technology solutions either do not get funded or don't show the results they hoped for - then they come running back to me to rely on the "oh, you mentioned we can save 5-20 percent through behaviour change...?"*
NATO policymaker

"*The problem lies in the sense of not believing the efficiency of behaviour change. Rather, it is a perception problem. The costs "appear" to outweigh the benefits and maybe it is sold too much as*

"behaviour" and could rather be sold as an aspect of a technology solution. Also, technocrats are funders and funding behaviour is too complex for them to address." Swiss researcher

"*It's important to work on practical implementation of behaviour change programmes and campaigns/projects. Behaviour change needs to be integrated in a wider, more technical programme. Results of behaviour change actions are not persistent over a short period. We need longer pilots and programmes to show persistence."* Finnish researcher

"*I think credibility of behaviour change programmes can be truly challenging. I am working on a pilot project and have always had this concern. How I can be sure of my results? What I have in my evaluation is based on interventions we used and how I can be sure it is not related to other causes? Meanwhile, when I go after fundraising for the projects, I need to explain the exact same thing to the funders."* UNIDO policymaker

"*Yes, especially to what concerns replicability. Project replicability of experiments in economics and the social sciences is hard. Long lasting efficiency of interventions (persistence) is often not shown. RCTs can help solving the credibility problem as they allow to be (in principle) replicated and provide a measure of cross-cultural variation."* EURAC researcher

"*There seems to be a large interest in behaviour change, but the societal structures don't follow. Funders such as national energy agencies don't have competencies to evaluate proposals from social scientists - this is clear when they still ask for numbers on how much kWh are reduced."* Swedish researcher

"*Within practice theory, I have struggled with designing policies and interventions and practical implementation in design. One argument I am often faced with from peers in the opposing camp is whether it all comes down to control and technological implementation and that this is just "talk."* UK researcher

"Sometimes it is difficult to prove to the government that behavioural measures actually work and that you can measure their effects." Italian policymaker

We also had an interesting discussion around **gender and class issues in behaviour change research**. Over 70 percent of the participants were female and many remarked how they seem to have to prove three times harder that their results are accurate than male colleagues (especially when the split is also between social scientists and engineers). On social justice issues, the hard-to-reach customer segments were mentioned most commonly, and several participants raised concerns about current behaviour change programmes being targeted too much at the white middle class.

"*I'm working on two projects on fuel poverty and the main issue (also highlighted by some third sector advisors) is that we are not able to communicate with them, to reach them and to build their trust. Moreover, my team encounters difficulties with social housing inhabitants (physical danger) when going to install energy meters in their homes (we now have to be accompanied by representatives of the social housing ownership)."* French researcher

"*From my experience we are rather talking about solutions that reach "scale" or solutions that reach an average consumer/household, since a lot of measures are affected by a self-selection bias."* Swiss researcher

"*I experienced problems with immigrants, especially because of language barriers. The respondents would have been happy to take part in the survey but they were not able to."* German researcher

"*My hard-to-reach customers are military personnel at an operational military base in West Africa! The only similarity I can draw is that it is mainly senior offices who I can draw access to / am directed to talk to - but actually the majority of the troops are the lower ranks and I want to talk to them!"* NATO policymaker

"*Hard-to-reach customers in South Korea are end users and office occupants. The questionnaires include user's occupancy time, break times, such as working patterns and particularly facility managers*

didn't want to participate in this research since they consider their reputation of their office and company too much." Korean researcher

"*Hard-to-reach customers in our case are people who are skeptical about the results of various research reports and don't have the will to change their behaviour. Also, customers who are poor, they live in bad conditions and have no money for simple energy measures. They feel even less privileged by the subsidies programmes*." Croatian policymaker

"*Coming from a country in the global south, it becomes highly difficult/questionable to forget lower social income groups regarding behaviour change. That is why in my opinion it is important to include urban planners, architects, and policy makers who are in the position to make a change where it is actually needed."* Egyptian researcher

"*1) in Greece, the economic crisis brought about a behaviour change in energy consumption 2) the power corporation, put a leaflet out years ago inside each bill, that reached out to all inhabitants of Greece."* Greek researcher

Many participants agreed that it was important to **focus on changing the behaviour of the *Behaviour Changers*.**

"*The government in Tajikistan does not rely on evidence to do policy making, including in the energy field. Therefore, starting at the local level (small scale) to do research and show them evidence in an effort to change their behaviour to look for advice before programming." Central Asian researcher*

"Yes, more communication and sharing of knowledge and experience among actors in the field as well as a more unified approach to monitor and evaluate." Italian policymaker

"*We need to be more open minded (we as scholars) and promote interdisciplinary discourse, we need to go beyond the professional ties that doesn't allow to see complementary other disciplines in informed solutions."* EURAC researcher

BECC conference, Washington D.C., October 7-10, 2018

At the BECC conference session, after a short overview of seven years of Task 24 findings, and the CEE involvement in Year 7, we convened a CEE sponsor panel who answered audience questions, and prepared questions around the topics of how many of their behavioural programmes could claim savings, how to evaluate them, and related regulatory barriers and solutions. The main problems discussed were **what behavioural savings programmes** were undertaken and **how they were evaluated**; and if there were any **regulatory barriers to implementing behaviour change programmes**.

How many behavioural programmes do you have and what savings are you claiming from them?

Jason Lenihan, Efficiency Vermont: We have five energy efficiency programmes focused on behaviour change; two are claiming savings, two are ongoing to determine whether savings are robust and viable (external) and also cost-effective (internal), and one more programme that we no longer do and aren't claiming savings.

Susan Norris, PG&E: We only use HERs, the Business Energy Reports didn't have sufficient savings to be able to tease out; we are trying NMEC via pilots, which is a combination of retrofit and attempted behavioural interventions. We look toward larger energy consuming residences in order to be able to see the savings at the meter level. We also do a number of programmes that are not considered behavioural but they are, including (1) schools programme k-12 and also pre-College and (2) Energy Management Training for building engineers that happens at the Pacific Energy Center. Also, we do a midstream retail products platform (RPP) programme to encourage retailers to carry more efficient products as opposed to less efficient products; I would argue that any market programme like RPP is a behaviour programme; you're not just changing the behaviour of the manufacturers but also the behaviour of the retailers. We have to focus too much on claiming savings versus what actually counts.

Arien Korteland, BC Hydro: BC Hydro has run energy efficiency programmes for more than 25 years, including Home Renovation Program, Electronics Program, and Low-Income Program. In addition, Team

Power Smart is part of BC Hydro's behavioural programme as a loyalty style programme. Residential customers can join and become a Team Power Smart member, which gives them exclusive perks, including member communications, special offers, member-only contests, and the opportunity to participate in an Energy Challenge and earn a monetary reward. We monitor a lot of metrics to keep track of programme health. Behaviour is an important topic for BC Hydro, and we try to take it very seriously. As an example, all programme managers are asked to see how they can embed behavioural insights into their programme designs.

Audience question: *How do your customers publicly commit to their energy use reduction?*

Arien: We have an online module that people can join and become a member and get a confirmation. Once they're a member they get access to a range of privileges, they receive a member dedicated newspaper. When we first started, we looked at other loyalty programmes like Weight Watchers, and after analysing these programmes they were more single-topic programmes, whereas the BC platform involves many different behaviours. The membership vehicle allows us to address a whole host of behaviours through that same umbrella. As a result, 12 months is a long time to keep people engaged, and we found that pride in the membership works to retain people's attention and motivation and retain them. We don't use a vendor, it's done completely in-house, we only claim energy from people who participate in challenges. Team Power Smart is cost-effective because we can claim savings from the entire membership group, some save energy without any incentives, which helps with the cost-effectiveness.

What's the methodology for claiming savings for behavioural programmes? What is the methodology used for claiming savings for other energy efficiency or DSM programmes?

Arien: We monitor more than 57 behaviours but found that looking at barriers for behaviour change is often more useful.

Jason: EVT has difficulty in getting customer data, which is especially challenging when data granularity is key. In 2017, we claimed savings for HERs, we also have a large SEM programme which includes 21 different members in cohorts including dairy processing, hospitals, and the biggest energy users (such as ski areas, Ben and Jerry's, universities, etc.). SEM persisted through R&D programme (tested via regression) and is now a resource acquisition programme - it's an 18-month campaign that begins with recruitment, then commitment workshops (where the plant manager or property manager who is the energy champion within the organisation will come with their decision-maker, talk about the programme, get an overview of the annual plan, and pledge publicly in front of the rest of the cohort, then we do kaizens and energy hunts). We are able to now claim savings from behavioural programmes as long as we do random assignment with Difference in Different technique or if we do a linear regression model. For one multi-family programme, anyone who saved energy in the programme was entered into a raffle to receive a free month's rent.

Susan: In California, behaviour is measured with RCTs, which are expensive and challenging, now we are testing NMEC programmes. We have rebate programmes that are easy to measure and deem the savings. We have a custom programme for our large commercial, industrial, and agricultural customers and one-by-one we show them how we can replace equipment and demonstrate savings as a result. Those are the approved methodologies that the regulators are comfortable with.

Audience: *What about public commitments? What about those who don't make a public commitment? Or those who do, but then don't comply?*

Arien: We did some public commitments a few years ago but then abandoned it, without comparing energy savings between the two groups. I would say that feedback is an important tool, but it cannot be the only behavioural tool. Sometimes we 'suffocate' our customers with information or feedback, while they just need a friendly nudge or reminder rather than a heavy dose of information. People don't need to be told everything there is to know about lighting when we just want them to switch off lights, because they know they should (although they don't necessarily always actually do it, e.g. because they forget). Each behaviour requires a different approach, sometimes information, sometimes a reminder.

What regulatory and/or organisational barriers, if any, do you perceive in terms of designing, implementing, and/or evaluating behavioural savings programmes?

Arien: The topic of persistence keeps coming up and we need to keep educating internally and externally.

Susan: Internally, it's more of a challenge of complying with regulatory requirements, now that it's been determined that 60 percent of programme implementation needs to be outsourced to third parties. Our regulatory requirements create some of these barriers given that we have to show influence and return on investment that is greater than 1:1. Our commission is safeguarding the public funds we use to fund these energy efficiency programmes, so they are safeguarding those dollars. We use the Total Resource Cost (TRC) calculation which has to be 1:1.25 at the portfolio level (so individual programmes can dip below that, but you can't have too many of those or they drag down your average).

Jason: We have an energy efficiency charge that's on the bill that the Vermont utilities send out, then we take those funds and deliver the energy savings-based on that charge. We started four years ago, and people didn't believe that behaviour wouldn't work in Vermont. The challenge is that we don't have the data and we don't have a direct link to the Vermont customers.

Audience: *How, if at all, do you incorporate NEBs?*

Arien: With our Team Power Smart programme we created a group of very loyal customers. From the participants of this programme we can relatively easily recruit for pilots, experiments, etc. There are likely other benefits as Team Power Smart members are more likely to recycle, etc.

Persistence: Who in the audience is studying it?

Kathy Kuntz: We disaggregate it, because some behaviours persist, whereas others don't, so it's important to distinguish (Cool Choices has some evaluation data on this).

Kevin: We have a project with Lumens. We can track energy usage and costs to low income customers over time and doing goal setting and back end measurement which would allow goal setting and then comparing how they do relative to that goal.

Hard-to-Reach Customers: What are we missing in terms of our definition?

- DTE Energy is doing work with deaf customers (Beth Fitzgerald at ESource and David Delind at DTE Energy)
- Renters and Multi-family (anonymous)
- Multi-family property manager are hard to get and engage (Doreen at PG&E)

IEA Demand Side Management Energy Technology Initiative

The Demand-Side Management (DSM) Energy Technology Initiative is one of more than 40 Co-operative Energy Technology Initiatives within the framework of the International Energy Agency (IEA). The Demand-Side Management (DSM) Energy Technology Initiative, which was initiated in 1993, deals with a variety of strategies to reduce energy demand. The following member countries and sponsors have been working to identify and promote opportunities for DSM:

Austria	Norway
Belgium	Spain
Finland	Sweden
India	Switzerland
Ireland	Canada
Italy	United Kingdom
Republic of Korea	United States
Netherlands	ECI (sponsor)
New Zealand	RAP (sponsor)

Programme Vision: Demand-side activities should be active elements and the first choice in all energy policy decisions designed to create more reliable and more sustainable energy systems
Programme Mission: Deliver to its stakeholders, materials that are readily applicable for them in crafting and implementing policies and measures. The Programme should also deliver technology and applications that either facilitate operations of energy systems or facilitate necessary market transformations

The DSM Energy Technology Initiative's work is organized into two clusters:
The **load shape cluster**, and
The **load level cluster**.

The 'load shape" cluster will include Tasks that seek to impact the shape of the load curve over very short (minutes-hours-day) to longer (days-week-season) time periods. Work within this cluster primarily increases the reliability of systems. The "load level" will include Tasks that seek to shift the load curve to lower demand levels or shift between loads from one energy system to another. Work within this cluster primarily targets the reduction of emissions.

A total of 24 projects or "Tasks" have been initiated since the beginning of the DSM Programme. The overall program is monitored by an Executive Committee consisting of representatives from each contracting party to the DSM Energy Technology Initiative. The leadership and management of the individual Tasks are the responsibility of Operating Agents.

These Tasks and their respective Operating Agents are:

Task 1 International Database on Demand-Side Management & Evaluation Guidebook on the Impact of DSM and EE for Kyoto's GHG Targets – Completed
Harry Vreuls, RVO, the Netherlands

Task 2 Communications Technologies for Demand-Side Management – Completed
Richard Formby, EA Technology, United Kingdom

Task 3 Cooperative Procurement of Innovative Technologies for Demand-Side Management – Completed
Hans Westling, Promandat AB, Sweden

Task 4 Development of Improved Methods for Integrating Demand-Side Management into Resource Planning – Completed
Grayson Heffner, EPRI, United States

Task 5 Techniques for Implementation of Demand-Side Management Technology in the Marketplace – Completed
Juan Comas, FECSA, Spain

Task 6 DSM and Energy Efficiency in Changing Electricity Business Environments – Completed
David Crossley, Energy Futures, Australia Pty. Ltd., Australia

Task 7 International Collaboration on Market Transformation – Completed
Verney Ryan, BRE, United Kingdom

Task 8 Demand-Side Bidding in a Competitive Electricity Market – Completed
Linda Hull, EA Technology Ltd, United Kingdom

Task 9 The Role of Municipalities in a Liberalised System – Completed
Martin Cahn, Energie Cites, France

Task 10 Performance Contracting – Completed
Hans Westling, Promandat AB, Sweden

Task 11 Time of Use Pricing and Energy Use for Demand Management Delivery- Completed
Richard Formby, EA Technology Ltd, United Kingdom

Task 12 Energy Standards - to be determined

Task 13 Demand Response Resources - Completed
Ross Malme, RETX, United States

Task 14 White Certificates – Completed
Antonio Capozza, CESI, Italy

Task 15 Network-Driven DSM - Completed
David Crossley, Energy Futures Australia Pty. Ltd, Australia

Task 16 Competitive Energy Services
Jan W. Bleyl, Graz Energy Agency, Austria / Seppo Silvonen/Pertti Koski, Motiva, Finland

Task 17 Integration of Demand Side Management, Distributed Generation, Renewable Energy Sources and Energy Storages
Seppo Kärkkäinen, Elektraflex Oy, Finland

Task 18 Demand Side Management and Climate Change - Completed
David Crossley, Energy Futures Australia Pty. Ltd, Australia

Task 19 Micro Demand Response and Energy Saving - Completed
Linda Hull, EA Technology Ltd, United Kingdom

Task 20 Branding of Energy Efficiency - Completed
Balawant Joshi, ABPS Infrastructure Private Limited, India

Task 21 Standardisation of Energy Savings Calculations - Completed
Harry Vreuls, SenterNovem, Netherlands

Task 22 Energy Efficiency Portfolio Standards - Completed
Balawant Joshi, ABPS Infrastructure Private Limited, India

Task 23 The Role of Customers in Delivering Effective Smart Grids - Completed
Linda Hull. EA Technology Ltd, United Kingdom

Task 24 Behaviour Change in DSM: Phase 1 - From theory to practice
Phase 2 – Helping the Behaviour Changers
Dr Sea Rotmann, SEA, New Zealand

Task 25 Business Models for a more Effective Market Uptake of DSM Energy Services
Ruth Mourik, DuneWorks, The Netherlands

For additional Information contact the DSM Executive Secretary, Anne Bengtson, E-mail: anne.bengtson@telia.com and visit the IEA DSM website: http://www.ieadsm.org